PHD

An investigation of mouse Basp1 and Wt1-AS, genes involved in the regulation of the Wilms' Tumour suppressor protein, Wt1

Allsop, Joanne

Award date:
2008

Awarding institution:
University of Bath

An investigation of mouse *Basp1* and *Wt1-AS*, genes involved in the regulation of the Wilms' Tumour suppressor protein, Wt1

Submitted by

Joanne Allsop

for the degree of Doctor of Philosophy

of the

University of Bath

Department of Biology and Biochemistry

2008

For my Family

Great things are not done by impulse, but by a series of small things brought together - Vincent van Gogh

Abstract

The Wilms' tumour suppressor protein, WT1, is a transcriptional regulator that plays an essential role in the development of several organ systems and has been found to be mutated in a number of diseases including Wilms' tumour of the kidney. Although the survival rates for this disease are high, other methods of treatment need to be sought as invasive treatments early in life can lead to debilitating illnesses during adulthood and can have severe short-term side effects. Although much has been found out about the genetics of Wilms' tumour in relation to the *WT1* gene, only 15% of all Wilms' tumours are considered to be due to an aberration in this transcriptional factor. For these reasons, further investigation into the genetics of this disease is needed.

Brain acidic soluble protein 1, BASP1, has been identified as a cosuppressor component that can modify regulation of the WT1 transcriptional activation domain by blocking its function. Within this thesis the murine expression pattern and genomic organisation of *Basp1* has been investigated. It was found that *Basp1* has an overlapping yet more extensive expression pattern than *Wt1*. A mouse knockout model of *Basp1* has also been generated. A breeding knockout line was not successfully established as all *Basp1* heterozygous knockout mice born died postnatally. The reasons for this postnatal lethality were not resolved.

WT1 is also known to have an overlapping antisense transcript, *WT1-AS* that produces an RNA that is not translated into protein. Several pieces of evidence point to *WT1* and *WT1-AS* being able to reciprocally regulate each other's expression by positive feedback. A mouse knockout model of the *Wt1-AS* gene has been successfully generated and phenotypic analysis has been presented. Histopathologically there were no striking abnormalities within the organs of these mice.

Table of Contents

List of Figures

List of Tables

Acknowledgments

Primarily my thanks go to my brilliant boss and supervisor Andrew Ward, who has allowed me the opportunity to carry out my studies and complete my thesis under his supervision. I would also like to thank Kim Moorwood who has taught me the practical side of science, has continually supported me throughout my time in the lab and is the best mentor one could hope for.

I would like to thank James Dutton, who always made time to look at my experimental problems despite being the busiest person I know and some of my other friends from the lab: Tia Smith, Treve Menheniott and Jeanette Müller, for encouraging me to embark upon a PhD.

For their expert technical help I would like to thank Iryna Withington for all things histological and the staff of 5 West Level 1, particularly Lesley Moore who put up with my nervousness in the early days. Also all of lab 0.76 past and present. No matter how crowded it gets in our tiny office, it has always been a great place to work and to share cake!

For their kind gifts of reagents and plasmids I would like to thank the following: William Skarnes (Sanger Institute, Cambridge) for the E14C4 ES cell line; Stefan Roberts (University of Manchester, Manchester) for the *Basp1* mouse cDNA probe; Nick Hastie (MRC, Edinburgh) for the WT1 BAC; Takeshi Yagi (NIPS, Japan) for the pMC1DTA plasmid; Marko Horb (McGill, Canada) for the LoxP-tPA plasmid; Philippe Soriano (FHCRC, USA) for the pGK neo tpA lox2 plasmid; Steven Sheardown (GSK, Harlow) for the pCAG-CRE and pSAβgeo plasmids; and Austin Smith (University of Cambridge, Cambridge) for the mGapdh plasmid.

Finally I would like to thank my family for their frequent encouragement and support, particularly my husband James who never doubted our mail would one day become homologous.

Abbreviations

17aa: 17 amino acids

APTS: 3-triethoxysilylpropylamine

ANOVA: analysis of variance

AIR: antisense IGF2R RNA

ARR: antisense regulatory region

AWT1: alternative WT1 transcript

BASP1: brain acid-soluble protein 1

BC1: brain cytoplasmic RNA 1

BIRPs: BASP1 immunologically related proteins

BSA: bovine serum albumin

bp: base pairs

BWS: Beckwith-Wiedemann syndrome

°C: degrees Celsius

CAG: chicken β-actin

CAP23: cortical cytoskeleton-associated protein 23

cDNA: complementary DNA

ChIP: chromatin immunoprecipitation assay

CKII: casein kinase II

cm: centimetre

CMV: cytomegalovirus

CNS: central nervous system

Cre: Cre recombinase

CTCF: CCCTC-binding factor

dCTP: deoxycytidine triphosphate

DDBJ: DNA Data Bank of Japan

DDS: Denys-Drash syndrome

DEPC: diethyl pyrocarbonate

DMD: differentially methylated domain

DMEM: Dulbecco's modified Eagle medium

DMR: differentially methylated region

DMSO: dimethyl sulfoxide

DNA: deoxyribonucleic acid

DNase: deoxyribonuclease

dNTPs: 2'-deoxynucleotide-5'-triphosphates

DPX: distyrene plasticiser in xylene

DTT: dithiothreitol

E: embryonic day

EDTA: ethylenediaminetetraacetic acid

EGR1: early growth response 1

EMBL: European Molecular Biology Laboratory

EMSAs: electrophoretic mobility shift assays

ES: embryonic stem

EST: expressed sequence tag

FACS: fluorescence-activated cell sorter

FCS: foetal calf serum

FS: Frasier syndrome

G418: geneticin selective antibiotic

GAP43: growth associated protein 43

GBM: glomerular basement membrane

GFP: green fluorescent protein

GST: glutathione-S-transferase

HA: haemagglutinin

hCG: human chorionic gonadotropin

HCl: hydrochloric acid

H&E: haematoxylin and eosin

HEK: human embryonic kidney

het: heterozygotes

HGMP: Human Genome Mapping Project

homo: homozygotes

HPRT: hypoxanthine phosphoribosyltransferase

HTF: human tubal fluid

IGF2: insulin-like growth factor 2

IP: intraperitoneal

IPTG: isopropyl-β-D-thiogalactoside

IU: international unit

IVF: *in vitro* fertilisation

kb: kilobase pairs

kDa: kilo daltons

KO: knockout

KTS: lysine, threonine, serine

L: litre

LacZ: β-galactosidase gene

LB: Luria-Bertani

LIF: leukaemia inhibitory factor

LOH: loss of heterozygosity

LOI: loss of imprinting

M: moles

MAR: matrix attachment region

MARCKS: myristoylated alanine-rich C kinase substrate

MCR: Mouse Conserved Region

MEM: minimum essential medium

mGFP: membrane bound green fluorescent protein

min/mins: minute/minutes

miRNA: micro RNA

mRNA: messenger ribonucleic acid

µF: microfarads

µg: microgram

µl: microlitre

µm: micrometers (micron)

µM: micromolar

mg: milligram

ml: millilitre

mm: millimetre

mM: millimolar

n: population size

$Na_3C_6O_7H_5 \cdot 2H_2O$: sodium citrate dihydrate

NAP22: neuronal axonal membrane protein 22

NCBI: National Centre for Biotechnology Information

ncRNA: non-coding RNA

neo^r: neomycin resistance gene

ng: nanogram

NH_3: ammonia

NLS: nuclear localization signal

O.C.T.: optimal cutting temperature compound

Ω: Ohms

P: probability value

P: passage number

PAGE: polyacrylamide gel electrophoresis

PAS: periodic acid-Schiff

PBS: phosphate buffered saline

PCR: polymerase chain reaction

PD: postnatal day

PFA: paraformaldehyde

pg: picogram

p*I*: isoelectric point

PIP: percentage identity plot

$PI(4,5)P_2$: phosphatidylinositol-4,5-bisphosphate

PKC: protein kinase C

PMSG: pregnant mares' serum gonadotrophin

PNS: peripheral nervous system

qPCR: quantitative PCR

RACE: rapid amplification of cDNA ends

rcf: relative centrifugal force (g)

RNA: ribonucleic acid

RNAi: RNA interference

RNase: ribonuclease

RO: reverse osmosis

RPA: ribonuclease protection assays

RT: reverse transcriptase

RT-PCR: reverse transcriptase polymerase chain reaction

SC: subcutaneous

SDS: sodium dodecyl sulphate

siRNA: small interfering RNA

smRNA: small modulatory RNA

snRNA: small nuclear RNA

snoRNA: small nucleolar RNA

sRNA: small RNA

SSC: saline sodium citrate

SV40: Simian vacuolating virus 40

TAE: Tris-acetate-EDTA

TE: Tris-EDTA

Tm: melting temperature

Tris: Tris [hydroxymethyl] amino methane

TTS: transcriptional terminator

Tukey's test: Tukey's multiple comparison post test

U: units of enzyme

UPD: uniparental disomy

UV: ultraviolet

V, kV: volts, kilovolts

WAGR: Wilms' tumour, Aniridia, Genitourinary abnormalities and mental retardation

wt: wildtypes

WT1: wilms' tumour 1

X-gal: 5'-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

YAC: yeast artificial chromosome

Introduction

1.1 Transcriptional control

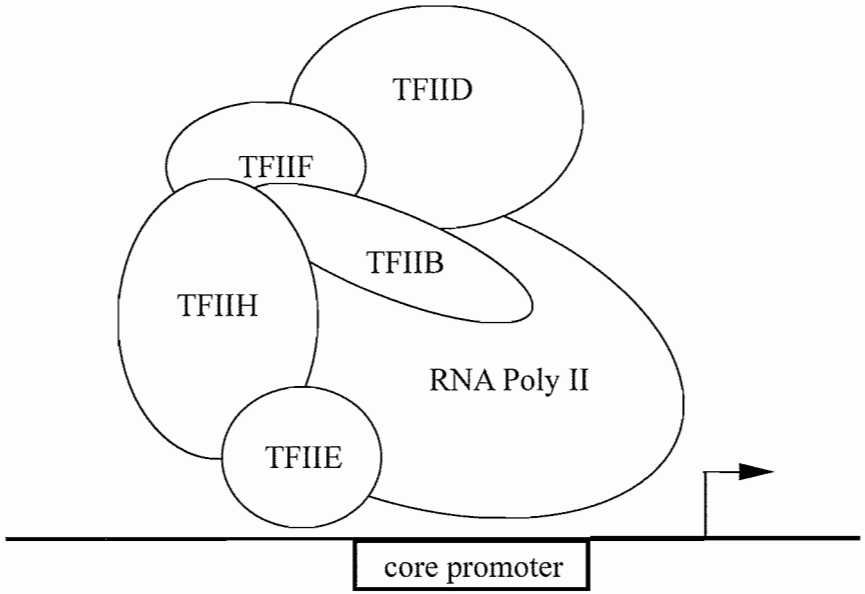
As is widely known, transcription factors are proteins that are able to alter the expression of genes either by upregulating or repressing their transcription. These proteins contain DNA binding domains that allow them to bind to specific DNA regions of genes they control. These DNA regions are termed response elements and can be either promoters or distal regulatory elements such as enhancers or silencers. The binding of transcription factors to these sites either prevents or allows subsequent transcription by blocking or aiding binding of RNA polymerase II and the other components of the basal transcriptional machinery, TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH to a core promoter adjacent to proximal promoter elements (reviews in (Maston et al., 2006) (Figure 1.1A). Transcription factors fall into several categories dependant upon the physical structure of their DNA binding domains. The most common types of transcription factors are the homeodomain proteins, helix-loop-helix, winged helix, zinc finger and leucine zipper proteins.

Transcription factors also possess activation and/or suppression domains to allow binding of cofactor proteins that do not contain DNA binding domains themselves but exert a regulatory effect on gene transcription by binding to these domains. Cosuppressors and coactivators are two types of such cofactors that bind to transcription factors altering their activity. Other cofactors that can assemble into RNA polymerases to take part in transcriptional control are chromatin remodelers, histone acetylases, histone deacetylases, kinases and methylases.

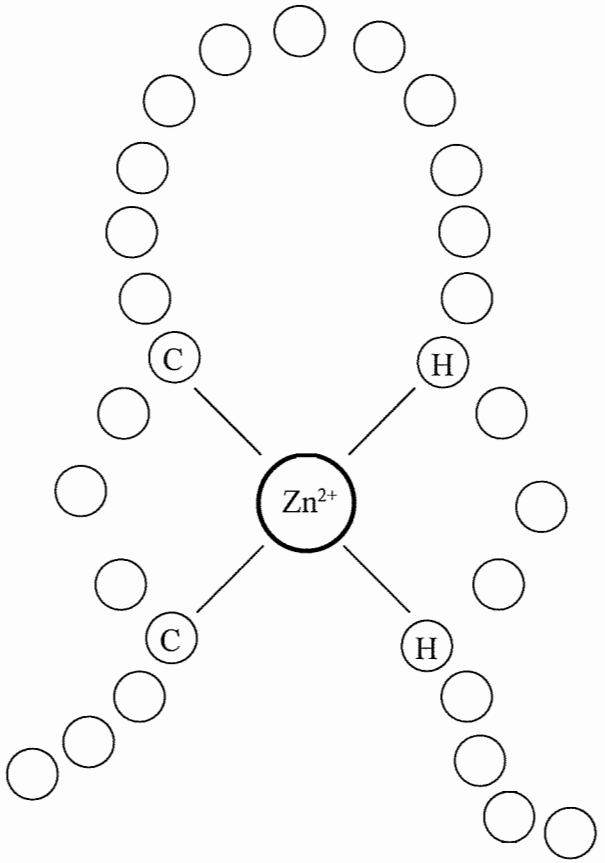
Due to their role in controlling the transcriptional activity of other genes, transcription factors are often involved in development and, in situations of aberrant expression, disease. Oncogenes for example can be transcription factors that are inappropriately activated, whereas tumour suppressor genes can be

Figure 1.1. (A) Schematic of the basal transcription machinery forming a preinitiation complex at a promoter. An arrow indicats the start of trancription. (B) Schematic of a C_2H_2 zinc finger motif. C, cysteine; H, histidine; O, amino acids; —, binding interactions with the zinc ion.

A


TFIID
TFIIF
TFIIB
TFIIH
RNA Poly II
TFIIE
core promoter

B


C
H
Zn2+
C
H

transcription factors that have been mutated thereby silencing their expression (reviewed in (Latchman, 1997) and references within). The Wilms' tumour 1 (*WT1*) gene is one such transcription factor; it is involved in development of the kidney and is also associated with the occurrence of Wilms' tumours of the kidney and other urogenital diseases. Surprisingly it has been shown to act as both an oncogene and tumour suppressor gene in different disease states.

WT1 is a zinc finger containing transcription factor. There are three main types of zinc finger domain, each containing a structural motif of cysteines and histidines surrounding a central Zn^{2+} ion: the C_2H_2, C_4, and C_6 zinc finger domains. The C_2H_2 domain (Figure 1.1B) is the most common and is known to fold into a secondary structure of two anti-parallel β-sheets and an α-helix. As well as binding to DNA, zinc fingers are able to bind to RNA and proteins (Gamsjaeger et al., 2007). WT1 and its cosuppressor protein Brain acid-soluble protein 1 (BASP1) will be discussed further in sections 1.4 and 1.5.

1.2 Mouse models of development and disease

The advent of embryology, culture of murine embryonic stem (ES) cells and molecular manipulation of mammalian genomes has allowed for sophisticated *in vivo* models of developmental growth processes and disease states to be examined. The mouse is an obvious choice as a model organism due to its high genomic sequence identity with humans, its relatively quick gestation time and the fact that it is the only species in which embryonic stem cells after culturing and manipulation have so far been successfully reintroduced into the embryo with a high efficiency.

The main strategies carried out to alter the gene dosage within the mouse are: mouse transgenics where additional copies of genes are added by pronuclear injection; mouse knockouts where copies of genes are inactivated; and mouse knockins where a gene is replaced with that of another. These strategies and the various techniques used are discussed below.

1.2.1 Mouse Transgenics

1.2.1.1 Pronuclear injection

This technique involves microinjecting a DNA solution of the cloned gene of interest, usually linked to a promoter of choice, into the pronuclei of fertilised oocytes before returning them to a recipient female. The DNA will be randomly yet stably integrated into the genome, often with tandem copies and even in alternate orientations although usually not at more than one locus. The integrated gene is then expressed, producing a gain of function phenotype. Due to the random sites of integration and copy number, expression levels may be difficult to control and the phenotype of such mice often differ between lines resulting in several lines having to be characterised to find phenotypes in common. Modifications to the cloned DNA can be carried out prior to injection in order to allow temporal and spatial expression of the cloned gene of interest.

1.2.2 Mouse knockouts

1.2.2.1 Targeted knockout

Targeting vectors containing a selectable marker and a reporter gene flanked by regions of homology to the gene to be disrupted can be introduced into ES cells and then screened for the presence of the transgene and its integration at the chosen locus by homologous recombination. These ES cells, once chosen *in vitro* can be reintroduced into a blastocyst by injection before being returned to a recipient female. Resulting pups that successfully uptake the ES cells will be chimeric and if the ES cells have contributed to the germline, some offspring of the chimeras will inherit one copy of the manipulated genome (reviewed in (Muller, 1999).

1.2.2.2 Conditional knockout

Like targeted knockouts, this technique involves disrupting a specific gene of interest. Unlike standard targeted knockouts however, conditional knockouts allow deletion of the gene of interest in a temporally and spatially controlled manner. This is advantageous for genes that are crucial for development and would therefore lead to prenatal or early death if absent during these stages, or

where the gene is expressed in a number of organs and only its role within one of these organs is of interest. Conditional knockouts are achieved using the same strategy of gene targeting in ES cells with the addition of recombination sites flanking the gene to be deleted; the introduction of a recombinase to allow deletion between these recombination sites can be achieved by crossing the resulting conditional knockout mice with transgenic mice expressing the recombinase under transcriptional control that will allow temporal expression or of a tissue specific promoter that will allow spatial control. The most common recombination system is the Cre-*loxP* system consisting of two 34bp *loxP* recombination sites and a Cre protein recombinase that is able to recognise and bind to these DNA sequences and cause a deletion between them. The most common method of transcriptional control of Cre is the tet-system made up of a transactivator protein that binds to a tet-operator sequence either in the presence or absence of tetracycline depending on the system, to cause expression of Cre (reviewed in (Muller, 1999).

1.2.2.3 Gene trapping

This is a technique used to disrupt genes at random with the ability to then trace where the insertion has taken place. Gene trapping can be achieved via promoter traps, gene traps and polyA traps. All three types of trap contain a selectable marker and a reporter gene and are designed to introduce a mutation into a gene. Promoter traps lack their own promoter and work by insertion into exons and are therefore transcribed using an endogenous gene promoter. Gene traps work by insertion into introns and, due to the presence of a splice acceptor, are transcribed using an endogenous promoter. PolyA traps contain a splice donor site and work by being spliced with exon sequences 3' of the integration site (cited in (Skarnes, 2000). Once the ES cells have been manipulated in culture, they are reintroduced into the mouse the same way as targeted mouse knockouts.

1.2.3 Mouse knockins

The replacement of one gene with another can be achieved by the modification of ES cells transfected with targeting vectors containing a selectable marker, a reporter gene and the gene of interest flanked by regions of homology to the gene

to be replaced. These ES cells containing the required genomic alterations can then be incorporated into the mouse using the same techniques as with knockout mice. As the replacement gene will be under the control of the original gene promoter, knockin mice provide a method to study gene pathways and gene redundancy (review in (Muller, 1999).

1.3 Non-coding RNAs

Non-coding RNAs (ncRNAs) are found in both prokaryotes and eukaryotes and their importance is becoming increasing apparent as more are being discovered in genome-wide screens. ncRNAs vary in size and function and are named accordingly: transcripts ~18-25 nucleotides in length belong to a group of either micro RNAs (miRNAs) or small interfering RNAs (siRNAs) involved in post-transcriptional gene silencing and RNA interference; transcripts ~20-300 nucleotides in length belong to a group known as small RNAs (sRNAs) which include small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs) and small modulatory RNAs (smRNAs) involved in modification of target RNAs, synthesis of telomeric DNA, chromatin structure dynamics, modulation of transcription and structural roles; larger transcripts that can be 300->10,000 nucleotides in length belong to a group known as medium or large RNAs involved in imprinting, DNA demethylation, X-inactivation, gene transcription and generation of the other classes of RNA (reviewed in (Costa, 2007; Storz, 2002). It is mainly the larger non-coding transcripts that are of interest to this thesis.

Within the mouse genome many potential ncRNAs were generated by the RIKEN Mouse Gene Encyclopaedia project that used an array of cell lines and tissues from many developmental stages to construct an extensive full-length cDNA library of 5'-capped and 3'-poly(A) tailed RNAs (cited in (Ravasi et al., 2006). Following sequencing and manual curation of a large set of these cDNAs (>60,000) by the FANTOM2 consortium, it was found that 47% lacked an apparent open reading frame and were therefore putative ncRNAs (cited in (Ravasi et al., 2006). A study to evaluate whether these transcripts have a

functional role was undertaken by Ravasi et al. (2006). Using a combination of RT-PCR, microarrays and Northern hybridization it was shown that many of these transcripts show tissue specific expression; this apparent regulation implies a biological function.

Many ncRNAs are divergently expressed and overlap protein-coding genes and are therefore termed antisense transcripts. Antisense transcripts can be spliced, polyadenylated and imprinted. Computational studies with the aim of identifying sense-antisense transcript pairs using EST databases (Yelin et al., 2003) and the FANTOM2 database (Kiyosawa et al., 2003) found approximately 2500 sense-antisense pairs and a further 900 pairs of divergently transcribed non-overlapping ncRNAs within the FANTOM2 database (Kiyosawa et al., 2003). Upon verification by strand-specific microarray analysis and Northern hybridization of these sense-antisense transcripts identified from the EST databases, it was found that approximately 60% were indeed divergently transcribed (Yelin et al., 2003), indicating that antisense ncRNAs are not as rare in the genome as was once thought and their abundance suggesting they may fulfil a functional role.

1.3.1 Functional validation of ncRNAs

As the genes coding for ncRNAs by their very nature do not code for protein products, validation of the roles of these genes has been more of a challenge than with protein coding genes. Three ways that have proven to be successful in this validation are RNAi knockdown, transcript knockouts and transcript truncations (Costa, 2007). An example of each of these methods is discussed below.

1.3.1.1 RNAi knockdown of *TUG1*

Taurine Upregulated Gene 1 (*TUG1*) is a 6.7kb ncRNA upregulated in murine developing retinal cells by taurine, a factor necessary for correct neural development (cited in (Young et al., 2005). Its function was ascertained by RNA interference (RNAi) knockdown assays whereby a small interfering RNA (siRNA) driven by a U6 promoter was coelectroporated with membrane bound green fluorescent protein (mGFP) driven by a CAG (chicken β-actin) promoter into rat retina at P0 and dissected at P25 in order to observe any physiological

cellular changes (Young et al., 2005). Photoreceptor rod cells of the retina appeared abnormal with malformed or absent outer segments and thin inner segments. The expression of several genes known to be important for rod development were also altered, indicating that *TUG1* was essential for the correct developmental regulation of these retinal cells (Young et al., 2005). The mechanism through which this ncRNA exerts its affect is unknown. It was postulated that the mechanism may be via globally altering chromatin configurations to aid the activation of other genes, although no experimental evidence was given.

1.3.1.2 Transcript knockout of *BC1*

Brain Cytoplasmic RNA 1 (*BC1*) is a ~150 nucleotide ncRNA hypothesised to be involved in higher brain functions and behaviour due to its expression patterns in nervous systems of rodents (cited in (Skryabin et al., 2003). In order to assess its function within the mouse, a gene targeting approach was carried out in ES cells followed by blastocyst injection to generate breeding chimeras whereby the entire transcript was knocked out by deletion of its promoter region (Skryabin et al., 2003). Homozygous knockout mice did not appear to exhibit any gross differences compared to wildtype siblings, however upon detailed behavioural analyses these mice were shown to exhibit behavioural differences consisting of reduced exploration and increased anxiety (Skryabin et al., 2003). *BC1* has therefore been shown to modulate brain activity and was postulated to confer a selective evolutionary advantage. The mechanism through which this ncRNA exerts its modulatory role is unknown.

1.3.1.3 Transcript truncation of *AIR*

Antisense Igf2r RNA (*Air*) is a paternally expressed 108kb ncRNA overlapping in an antisense orientation, the maternally expressed, paternally methylated, *Igf2r* protein coding gene. The *Air* promoter begins in intron 2 of *Igf2r* within an imprint control element for these and several other genes of an imprinted gene cluster. To determine the role of *Air*, a manipulation of the mouse genome was performed whereby the functional full-length transcript was truncated by 96% during transcription. This was achieved by a gene targeting approach in ES cells

followed by blastocyst injection to generate breeding chimeras whereby a polyadenylation signal was inserted immediately downstream of the imprint control element containing its promoter so as not to affect this control region (Sleutels et al., 2002). When the targeted allele was inherited from the maternal allele the phenotype of the mice was indistinguishable from that of wildtype siblings. When the targeted allele was inherited from the paternal allele or mice were homozygous for the targeted allele mice exhibited a 15% reduction in birth weight compared to wildtype siblings; in addition, paternal methylation of *Igf2r* was lost, and paternal silencing of two other genes within the same imprinted gene cluster divergently expressed in relation to *Air* was also lost. Another experiment showing the importance of the *Air* transcript involved crossing male mice with the *Air* truncation, with female mice containing a 3cM chromosomal deletion spanning the *Igf2r* locus. Usually, maternal inheritance of the *Igf2r* deletion is lethal with pups dying between E15.5-17.5; these double heterozygotes however were of normal size, viable and fertile indicating that the lethal phenotype was rescued by the paternal truncation of *Air* (Sleutels et al., 2002). *Air* has therefore been shown to be required for *cis*-acting silencing of flanking imprinted protein coding genes (Sleutels et al., 2002). *Air* was subsequently found to be a capped and polyadenylated, unspliced transcript transcribed by RNA polymerase II, retained within the nucleus and highly unstable (Seidl et al., 2006). The mechanism through which this ncRNA causes gene silencing was postulated to be due to the event of transcription through a *cis*-acting domain regulatory element within the locus, rather than its physical presence; this transcriptional interference resulting in the RNA polymerase II at the *Air* promoter continually moving across the *Igf2r* promoter.

1.3.2 ncRNAs at imprinted loci

Imprinting is often tissue specific and developmentally controlled. Many ncRNAs are associated with imprinting clusters such as the one described above containing *AIR* and may therefore be mechanistically involved in the role of complex gene control. On the other hand, the association of ncRNAs at imprinted clusters could be due to the so far limited effort of searching for ncRNAs at other loci. Other imprinted ncRNAs that form part of an imprinting locus include: *H19*

at the *IGF2* locus, *NESPAS* and *1A* at the *GNAS* locus, *XIST* and *TSIX* at the *XIST* locus, as well as, *KCNQ1ot1* (*Lit1*), *Copg2AS2*, *Dio3AS*, *Ube3A-AS* (reviewed in (O'Neill, 2005; Tufarelli, 2006).

H19 at the *IGF2* locus was the first ncRNA found to be associated with an imprinted gene cluster. *H19* is a maternally expressed, spliced, polyadenylated transcript of unknown function. It is not thought to be involved in imprinting control of the locus, as shown by the normal phenotype of *H19* knockout mice and the identification of an imprinted control element of both *Igf2* and *H19* termed the differentially methylated domain (DMD) (reviewed in (O'Neill, 2005). Due to the high sequence identity of *H19* between the mouse and human genomes (77%) however, it was postulated to be functional.

The *XIST* locus has been intensely studied due to its involvement in X-chromosome inactivation in female mammals. The *XIST* transcript is only transcribed from the chromosome that will later become inactive; this was found to be due to *cis* effects of the XIST RNA physically coating the inactivated X-chromosome. XIST is therefore a non-coding gene silencer. In mouse extraembryonic tissue, the overlapping maternally expressed non-coding *Tsix* antisense transcript is thought to repress expression of the *Xist* transcript, directing paternal chromosomal inactivation. In human extraembryonic tissue however chromosomal inactivation is random and *TSIX* may therefore have another role other than an antagonist of *XIST* (reviewed in (O'Neill, 2005). In the embryo proper X-chromosome inactivation is not imprinted within the mouse or human. Experiments within the mouse genome whereby *Tsix* transcripts were truncated prior to an overlap with *Xist* showed that *Tsix* was no longer able to block *Xist* expression. When a reverse orientation *Tsix* cDNA was then knocked-in to this locus to allow transcription of the full length *Tsix* to observe the *trans* effect of this transcript, repression of the *Xist* transcript was not restored. This epigenetic process of gene silencing therefore needs the antisense transcription of *Tsix* through the *Xist* locus rather than the physical presence of the *Tsix* transcript, a situation parallel to the transcription of the *Air* transcript antisense to *Igf2r* (reviewed in (O'Neill, 2005) and references within).

1.3.3 ncRNAs at the WT1 locus

At the WT1 locus there are overlapping antisense transcripts that are imprinted, at least in the human genome. These antisense transcripts are the subject of part of this thesis and will be discussed further in section 1.4.10.

1.4 Wilms' Tumour and the *WT1* gene

1.4.1 Wilms' Tumour Disease

Wilms' Tumour or Nephroblastoma is a common paediatric tumour of the kidneys that affects approximately 1 in 10,000 children below the age of 5 years. The survival rate for this disease is approximately 85% with a combination of chemotherapy, radiotherapy and surgery being the means of treatment (Brown and Malik, 2001). Although the survival rates for this disease are high, other methods of treatment need to be sought, as invasive treatments early in life can lead to debilitating illnesses during adulthood and can also have severe short-term side effects. With these points in mind, an investigation into the genetics of this disease seems to be the way forward.

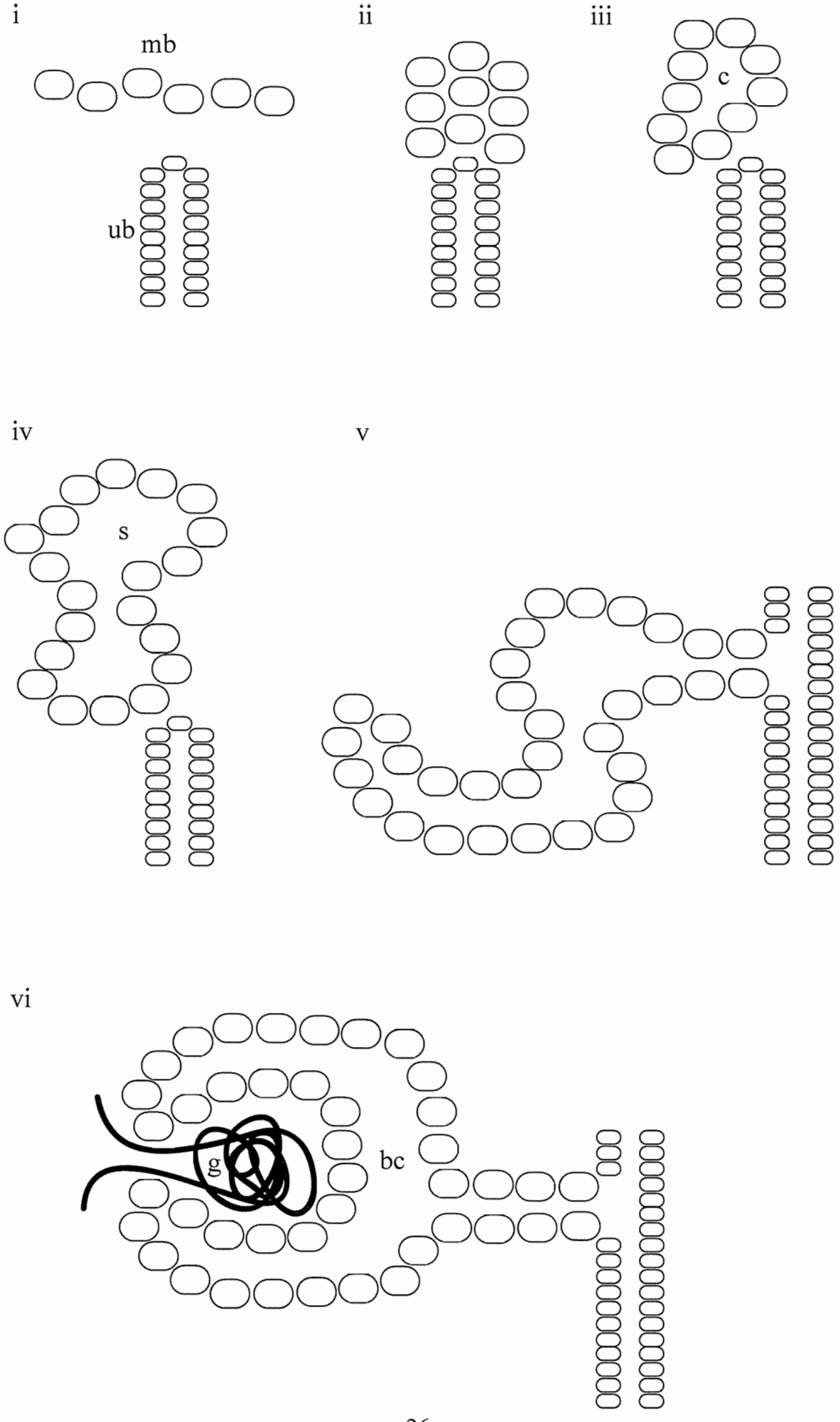
1.4.2 Kidney development

In order to understand the process that occurs during kidney malformation it is essential to have an appreciation of the normal developmental patterns of this organ. During higher vertebrate embryogenesis three pairs of increasingly sophisticated excretory organs develop: the pronephros, mesonephros and metanephros. The pronephros is formed of simple nephron structures attached to the pronephric duct that do not have integrated tubules and glomeruli/glomera; these structures are instead situated either side of a cavity region called the nephrocoel. In some lower vertebrates such as the lamprey primitive fish, the pronephros is the functioning excretory system: waste products are filtered from the vascularized glomeruli into the nephrocoel where they are then moved into the tubules by the cilia action of structures called nephrostomes allowing uptake of nutrients back into the blood system and removal of waste via the pronephric duct. In higher vertebrates this structure, which forms around four weeks of

gestation in humans, regresses as embryogenesis progresses (Martini et al., 2002; Vize et al., 1997). The mesonephros is formed when the growth of the pronephric duct causes induction of the intermediate mesoderm into mesonephric tubules, occurring at the end of four weeks of gestation in humans. The nephron structures of the mesonephros unlike the pronephros have integrated tubules and glomeruli and in some vertebrates, including mammals, the glomeruli are surrounded by Bowman's capsules. The mesonephric tubules fuse with the pronephric duct to form the mesonephric duct (Wolffian duct) thus completing the development of the functioning excretory system of some vertebrates such as fish and certain amphibians. In higher vertebrates, the mesonephros is replaced by the metanephros or definitive kidney, occurring at around five weeks of gestation in humans (Martini et al., 2002; Vize et al., 1997).

The definitive kidney begins to form in parallel with the gonads and has hence been termed the urogenital system. The initial signalling in the development of the kidneys occurs when a column of cells called the ureteric bud, that is formed from an outgrowth of the Wolffian duct, migrates into the metanephric blastema (Figure 1.2). The ureteric bud induces condensation of this surrounding metanephric mesenchyme followed by a burst of proliferation, and reciprocally the metanephric mesenchyme induces the ureteric bud to branch. The condensation of mesenchymal cells causes a cavity, termed the renal vesicle, to form around the tip of the ureteric bud. The majority of cells comprising the renal vesicle then undergo differentiation, often termed the mesenchymal-to-epithelial transition, during which the cells of the vesicle form transitory structures called comma-shaped, followed by S-shaped bodies. These transitory structures then fuse with the branching ureteric bud and undergo elongation to form nephric structures: the Bowman's capsule, proximal and distal tubules and the loop of Henle. During these stages the branching ureteric bud forms the collecting duct system and ureter. The rest of the metanephric mesenchymal cells do not develop into nephrons, but form the connective tissue of the kidney or undergo apoptosis (reviewed in (Brown and Malik, 2001; Hastie, 1994; Reddy and Licht, 1996)).

Figure 1.2. Schematic diagram of the stages of kidney development. i, the ureteric bud migrates into the metanephric blastema. ii, the ureteric bud induces condensation of the surrounding metanephric mesenchyme followed by a burst of proliferation. iii, the renal vesicle starts to form with the transitory comma-shaped body structure. iv, the transitory S-shaped body structure forms. v, the transitory structure fuses with the branching ureteric bud and undergoes elongation to form nephric structures. vi, the Bowman's capsule is formed surrounding the glomerulus. ub, ureteric bud; mb, metanephric blastema; c, comma-shaped body; s, S-shaped body; bc, Bowman's capsule; g, glomerulus. Figure adapted from Reddy and Licht (1996).


i
mb
ub
ii
iii
c
iv
s
v
vi
g
bc

1.4.3 Kidney function

Nephrons are the functional units of the kidney that filter the blood of waste products. The blood enters the glomerular capillaries of the nephron via the afferent arteriole where water and solutes under high pressure are excreted from these capillaries into the Bowman's capsule. The excretion of large molecules from the blood such as proteins is prevented by the glomerular basement membrane (GBM) and the podocyte cells that form a meshwork connected by slit diaphragms surrounding the glomerular capillaries and GBM. The filtrate then enters the proximal convoluted tubule where water and other vital nutrients are reabsorbed into the blood via the efferent arterioles before being passed along the loop of Henle consisting of a descending and an ascending limb. The descending limb carries the filtrate from the cortex to the medulla allowing further osmotic escape of water into the efferent arterioles due to the surrounding hypotonic environment; the ascending limb carries the filtrate back to the cortex and into the distal convoluted tubule, pumping out salts whilst remaining impermeable to water. Once the filtrate has reached the distal convoluted tubule it is actively transported into collecting tubules where it flows to the collecting ducts and is released into the renal calyces and renal pelvis before being passed through the ureter to the bladder as urine (reviewed in Vize et al., 1997).

1.4.4 Wilms' tumours

Wilms' tumours can be sporadic or hereditary and occur unilaterally or bilaterally with the presence of one or multiple tumours. Hereditary cases of the disease, known to be inherited autosomal dominantly, only account for approximately 1% of Wilms' tumours (Matsunaga, 1981). Wilms' tumours also occur in association with syndromes including the 'Wilms' tumour, Aniridia, Genitourinary abnormalities and mental retardation' (WAGR) syndrome, Denys-Drash syndrome (DDS) and Beckwith-Wiedemann syndrome (BWS) amongst others. WAGR syndrome as the title suggests is a combination of tumours of the kidney; aniridia, an eye malformation where the iris is absent; genital abnormality and mental retardation, with Wilms' tumours appearing in 30% of patients. DDS involves genitourinary defects whilst BWS involves overgrowth of organs with

the incidence of Wilms' tumours being 30% and 5% respectively (reviewed in (Rivera and Haber, 2005).

Wilms' tumours are heterogeneous, specifically triphasic, displaying characteristics of blastemal, epithelial and stromal tissue types, and are therefore theorized to be derived from a pluripotent cell type (reviewed in (Rivera and Haber, 2005). Indeed, Wilms' tumours are thought to arise from metanephric mesenchymal stem cells of the kidney blastema that fail to enter into the epithelial transition stage and instead continue to proliferate. As a result of this lack of differentiation, nephrons fail to form and the undifferentiated stem cells form areas of premalignant embryonic tissue termed nephrogenic rests. Although it is possible for these nephrogenic rests to regress, and the majority do, others go on to become neoplastic rests that can become malignant and form Wilms' tumours (reviewed in (Hastie, 1994; Malik et al., 2001). The histology of Wilms' tumours supports this immature kidney phenotype with sections of the blastema resembling the transitory structures that occur during the mesenchymal-to-epithelial transition (Rivera and Haber, 2005). Wilms' tumours therefore differ from other paediatric tumours of the kidney such as clear cell sarcoma in that they arise from foetal kidney development gone awry.

Wilms' tumours can appear with or without allele loss at the 11p13 locus. Interestingly, where allele loss occurs it seems to always occur by the non-random loss of the maternal allele, suggesting a role for genomic imprinting in Wilms' tumour progression (Schroeder et al., 1987).

1.4.5 Testis Development

The urogenital ridge is made up of the pronephros, the mesonephros and the metanephros. Gonadal development begins with a thickening of the genital ridge epithelium within the mesonephric region giving rise to the sex cords. The germ cells migrate into this region and are surrounded by the sex cords. The sex cords continue to develop and fuse to form the rete testis and the seminiferous cords that contain the Sertoli cells, structural cells that will later provide nutritive support to the developing gamete cells. A layer of connective tissue separates

away from the seminiferous cords to form the surrounding tunica albuginea interspersed by mesenchymal cells. Within the testis, the interspersing mesenchymal cells form Leydig cells which produce testosterone in order to maintain the mesonephric duct (Wolffian duct). The Sertoli cells of the seminiferous cords meanwhile secrete anti-Müllerian hormone to inhibit the development of the female reproductive Müllerian ducts. The tubules of the mesonephros join up to the rete testis to form the efferent ducts that link the testis to the Wolffian duct. The Wolffian duct then differentiates to form the structures of the epididymis and vas deferens (Gilbert, 2000).

During puberty the seminiferous cords develop to become the seminiferous tubules containing the germ cells that will form mature gametes through the process of spermatogenesis. During this process, diploid spermatogonium mitotically divide to produce primary spermatocytes which then undergo a first meiotic division to produce haploid secondary spermatocytes. These secondary spermatocytes undergo a second meiotic division to produce haploid spermatids which mature to become the mature gamete cells, sperm.

1.4.6 The *WT1* gene

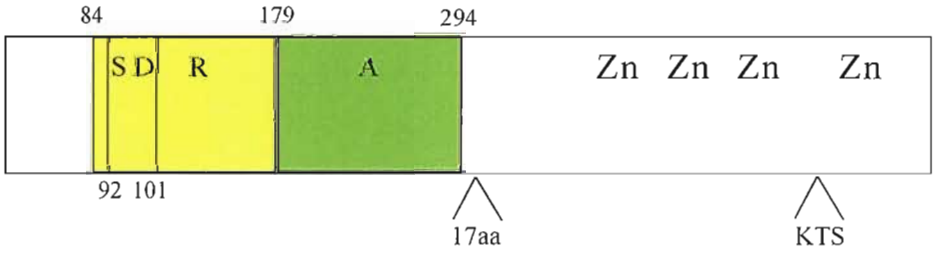
By virtue of pedigree studies of children with the WAGR syndrome and LOH in Wilms' tumours, the Wilms' tumour locus was mapped to chromosome position 11p13 (Gessler et al., 1990) and was later designated as the *WT1* locus. *WT1* spans approximately 50kb of the genome, it is comprised of 10 exons, has three transcriptional start sites generating RNA transcripts of approximately 3kb in length and has a GC-rich upstream regulatory region with a promoter lacking both TATA and CCATT boxes (reviewed in (Reddy and Licht, 1996). The RNA transcripts were found to be subject to alternative splicing generating protein products of 52-54kDa predicted to contain four C-terminal zinc finger motifs of the Krüppel C2-H2 class characteristic of transcription factors, and an N-terminal proline-glutamine-rich region (Gessler et al., 1992). There are over 28 isoforms of WT1 generated by alternative splicing, RNA editing, an alternative translational start site, and an alternative promoter and leader exon (reviewed in (Hohenstein and Hastie, 2006). The most common splicing events cause an

insertion of 17 amino acids (17aa) upstream of the zinc fingers encoded by exon 5; and an insertion of 3 amino acids, lysine, threonine and serine (KTS) between zinc fingers 3 and 4, encoded by an alternative splice site of exon 9. These main isoforms are thought to be present in fixed ratios of 8.3 : 3.8 : 2.5 : 1 (17aa+/KTS+ : 17aa-/KTS+ : 17aa+/KTS- : 17aa-/KTS-) in mammals (Haber et al., 1991); although the KTS isoform is conserved amongst vertebrates, the 17aa splice is mammal specific (reviewed in (Scharnhorst et al., 2001). The KTS isoforms have distinct yet overlapping DNA binding activities and occupy distinct sub-nuclear compartments suggestive of independent roles for each of these proteins (Larsson et al., 1995). RNA editing occurs by the replacement of a leucine residue for a proline residue caused by a U to C transition within exon 6 (Sharma et al., 1994). Alternative translational initiation occurs from a second in-frame AUG (Scharnhorst et al., 1999) and a mammal-specific upstream CUG in-frame with the main AUG start site (Bruening and Pelletier, 1996). The transcript derived from an alternative promoter and leader exon (alternate exon 1 or exon 1a) has been named (*Alternative WT1 transcript*) *AWT1*. The AWT1 protein is 33kDa in size and exhibits the 17aa and KTS splices also present in the other WT1 isoforms. Both the promoter region and exon 1a are located in intron 1 of *WT1* (Dallosso et al., 2004). This transcript was initially identified in a human testis cDNA library and subsequently in human foetal kidney and mouse kidney by RT-PCR (Dallosso et al., 2004). In the mouse the *Wt1* locus is located on chromosome 2 and the same protein isoforms are produced.

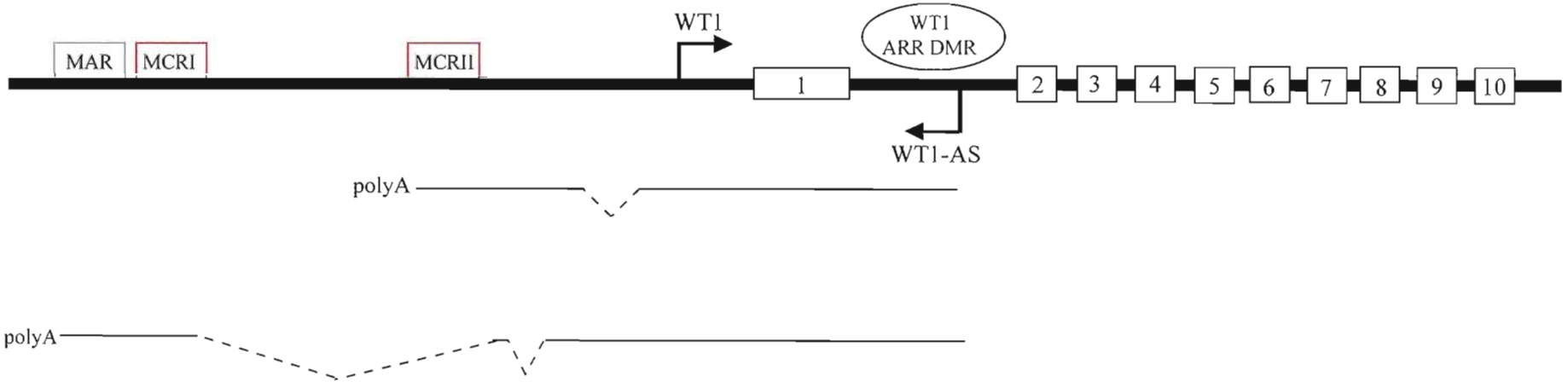
As well as four zinc fingers, WT1 has both a transcriptional activation domain and a transcriptional repression domain (Figure 1.3A). These domains were characterised by transient cotransfection assays in a murine fibroblast cell line (NIH3T3) of WT1 N-terminal deletion mutants and reporter plasmids containing WT1 binding sites either side of a minimal promoter of the *Platelet-derived growth factor A-chain* (*PDGF A-chain*) gene (a gene transcriptionally suppressed by WT1) (Wang et al., 1993). When an expression construct containing full length WT1 was cotransfected with a reporter plasmid with WT1 binding sites placed upstream and downstream of the *PDGF A-chain* transcriptional start site this served to repress the transcriptional activity of the reporter. When a WT1 N-terminal deletion mutant lacking amino acids 1-84 was cotransfected instead of

Figure 1.3. (A) Schematic of the WT1 protein. Positions of the Suppression domain (SD), Repression domain (R), Activation domain (A) and the four zinc fingers (Zn) are shown. (Figure adapted from Wagner and Roberts (2004)). (B) Schematic of the *WT1* and *WT1-AS* genes. The gene locus is represented by the thick black horizontal line, the 10 exons of *WT1* in black open boxes, the *WT1* ARR DMR between *WT1* exon 1 and intron 1 in the circle, the relative positions of the *WT1* and *WT1-AS* promoters shown as a right-angled arrows, the conserved regions MCRI and MCRII in open red boxes, the matrix attachment region (MAR) in an open grey box and two spliced *WT1-AS* transcripts underneath the gene locus. The shorter antisense transcript shows the splice identified by Gessler and Bruns (1993) and the exon 1 overlapping region originally reported by Campbell et al. (1994). The longer antisense transcript shows the position of two splices identified by Dallosso et al. (2007). This longer transcript lacks transcription of the MCRII region.

A


84
179
294
S D
R
A
Zn Zn Zn Zn
92 101
17aa
KTS

B


MAR
MCRI
MCRII
WT1
WT1
ARR DMR
1
2
3
4
5
6
7
8
9
10
WT1-AS
polyA
polyA

full length WT1 the same repression was observed. However, when a WT1 N-terminal deletion mutant lacking amino acids 1-179 was cotransfected rather than full length WT1, repression did not occur (Wang et al., 1993). This array of experiments served to narrow down the transcriptionally active sites of WT1 to a repression domain contained within amino acids 84-179 and an activation domain contained within amino acids 179-294 (Figure 1.3A). Interestingly, the WT1 isoform AWT1 lacks the repression domain (Dallosso et al., 2004).

Overlapping the repression domain of WT1 there is another more specific region termed the suppression domain that can suppress its own activation domain (Figure 1.3A). GAL4-fusion proteins of the WT1 repression domain (GAL4-R), WT1 activation domain (GAL4-A) and both domains (GAL4-RA) driven by a cytomegalovirus (CMV) promoter were cotransfected into a human embryonic kidney (HEK) cell line (293 HEK cells) along with a reporter construct containing GAL4 binding sites (McKay et al., 1999). Compared to GAL4-R and GAL4-RA, GAL4-A exhibited a significant transcriptional activation of the reporter construct, indicating that in the presence of the repression domain the transcriptional activity of the activation domain is repressed. The physical locality of this suppression domain was defined by cotransfection of WT1 deletion mutants fused to the activation domain together with a reporter construct in 293 HEK cells. The suppression domain was initially found to be contained within amino acids 71-101 (McKay et al., 1999) and subsequently narrowed down to amino acids 92-101 (Carpenter et al., 2004).

As well as Wilms' tumours and WAGR syndrome, WT1 mutations are also associated with DDS and Frasier Syndrome (FS). Both syndromes involve genitourinary defects, however, FS patients rarely exhibit Wilms' tumours (compared to 30% of DDS patients). Deletion of a single copy of *WT1* is not sufficient for the development of WAGR syndrome or Wilms' tumours and is therefore not due to WT1 haploinsufficiency. Children who inherit a *WT1* mutation from a parent carrier, however, have a high risk of developing Wilms' tumours. This is because only one *WT1* mutation is required in the second *WT1* allele during kidney development for the emergence of Wilms' tumours to be possible. This genetic situation is described as Knudson's 'two-hit model' of

tumourigenesis. The first occurrence of this phenomenon was reported by Knudson (1971) in association with the first discovered tumour suppressor gene, *Retinoblastoma 1* (*RB1*). *RB1* is the only gene to be associated with retinoblastoma, a paediatric tumour of the retina. When *RB1* is inherited from an affected parent there is a high probability that children will go on to develop retinoblastoma if the second copy of the gene is lost or mutated. DDS is thought to arise due to dominant-negative actions of mutant WT1 protein caused by point mutations in the zinc finger domain, whereas FS is thought to arise due to a shift in the ratio of WT1 KTS- and WT1 KTS+ isoforms due to point mutations in the KTS splice donor site at the exon 9 3' end (reviewed in (Hastie, 1994; Rivera and Haber, 2005). Although BWS patients exhibit Wilms' tumours, chromosomal rearrangements have been mapped to 11p15 and are therefore not linked to WT1. Likewise, familial cases of Wilms' tumours do not map to the *WT1* locus (reviewed in (Hastie, 1994). These observations that the appearance of Wilms' tumours are not caused by a loss of function at a single locus in all hereditary or sporadic cases came as a disappointment to many in the field as it did not follow the same simple genetics of retinoblastoma. The conclusion being that Wilms' tumours do not always follow Knudson's 'two-hit model' of tumourigenesis.

An interesting observation is the non-random loss of the 11p allele region in Wilms' tumours, with the paternal allele being retained (cited in (Little et al., 1992). This observation led to the suggestion of the involvement of an imprinted gene at this locus. *WT1* however has been shown to be equivalently expressed from both the paternal and maternal allele in normal foetal kidney and Wilms' tumours (Little et al., 1992). Interestingly, *WT1* appears to be monoallelically expressed in some placental and foetal brain tissue with *WT1* expression arising mainly or exclusively from the maternal allele in these imprinted placental tissues (Jinno et al., 1994). *WT1* imprinting has therefore been proposed to be tissue specific and genetically polymorphic within the human population (Jinno et al., 1994). Although *WT1* is not imprinted in the kidney, recent evidence has found imprinting of both *AWT1* and a *WT1* antisense transcript within this organ (see section 1.4.13).

1.4.7 Sites of expression

Much has been done to elucidate the sites of *WT1* expression in both foetal and adult tissues, many experiments overlapping in their findings. Pritchard-Jones et al. (1990), using human foetal kidney *WT1* cDNAs, showed by Northern hybridization analysis of human foetal tissue that *WT1* was expressed in kidney, spleen, testis, ovary and weakly in the brain. By *in situ* mRNA hybridization they showed that in human foetal kidney, *WT1* was specifically expressed in the condensing metanephric blastema, metanephric vesicles and the renal corpuscle, particularly the developing podocytes of the glomerular epithelium; in human foetal testis expression was seen in the developing gonadal ridge, particularly over the sex cords (Pritchard-Jones et al., 1990). These sites of RNA expression in the genitourinary system have been verified, once again by *in situ* mRNA hybridization, and further, the sites of protein expression have been shown by immunohistochemistry with anti-WT1 monoclonal antibodies (Mundlos et al., 1993). In the developing kidney, protein was expressed in the glomerulus, specifically the podocytes, but not in the metanephric blastema or renal vesicles. In the developing testis, protein was expressed over the tunica albuginea connective tissue, the destined seminiferous tubules and rete testis, progressively restricting to the Sertoli cells of the seminiferous tubules and the rete testis. They also showed that whereas RNA levels in the glomerulus were highest in foetal stages (and could only be detected by RT-PCR in the adult), protein levels in the same structures were highest in the adult (Mundlos et al., 1993). High levels of WT1 were also found in both human and murine bone marrow using immunological FACS (fluorescence-activated cell sorter) sorting techniques to select for bone marrow precursor cells (Fraizer et al., 1995).

Pelletier et al. (1991) and Park et al. (1993) used the same techniques of Northern and *in situ* mRNA hybridization with murine *Wt1* cDNAs and showed expression in the myometrium of the uterus and the non-germ cells of the gonads, specifically the Sertoli cells of the testis and the granulosa and epithelial cells of the ovaries (Pelletier et al., 1991), the mesothelial linings of the pericardial space surrounding the heart and diaphragm, the pleural space surrounding the lungs and the peritoneal space surrounding the liver and small intestines, diffuse expression within the spleen later identified as the stromal cells and strong expression in the

surrounding mesothelial capsule of both the spleen and thymus (Park et al., 1993). Northern hybridization analysis of murine tissue has also shown *Wt1* to be expressed in the adult, specifically in the kidney and spleen, and at lower levels in the heart, lung, thymus, testis and ovary (Buckler et al., 1991). A detailed examination of the temporal and spatial expression of *Wt1* in the developing human and mouse embryos by *in situ* mRNA hybridization was carried out by Armstrong et al. (1993). At developmental stages where human tissue samples were available the expression patterns correlated with those seen in the mouse. *Wt1* expression was shown to begin at early mouse embryonic day (E) 9 where it was restricted to the lateral lining of the coelomic cavity and its surrounding epithelium. At late E9 it was expressed in the mesothelium surrounding the cavities of the heart and gut, and over the urogenital ridge, particularly the mesonephros. Expression persisted in these regions during development at progressively restricted sites. Additionally, expression could be seen in the mesothelial linings of organs contained within the coelomic cavity at E10; weak expression was seen in the spinal cord from E11.5 and the roof of the fourth ventricle of the brain at E15. By E19 expression was only seen in the kidney and gonads where it persisted in adult tissues (Armstrong et al., 1993). In addition, RT-PCR analysis showed *Wt1* to be expressed in the eye and tongue in both foetal and adult tissues. Unlike observations of Buckler et al. (1991) however, no expression was observed in the thymus.

To further elucidate the expression patterns of *Wt1* during murine development Rackley et al. (1993) undertook a study looking at both mRNA and protein expression using *in situ* mRNA hybridization and immunohistochemistry techniques. The *in situ* pattern correlated with the findings of Armstrong et al. (1993), and added the abdominal wall muscle and uterus of embryonic and pregnant females to the list of expression sites. Antibody staining was predominantly nuclear, and was observed at E15.5 in the kidney, testis, uterus and spinal cord (Rackley et al., 1993). Within the kidney, antibody staining was present in metanephric blastemal cells, the renal cortex, and the early condensations and podocyte cells of the metanephric vesicles. Using the same techniques, mRNA and protein expression was found in the olfactory epithelium between E9.5-E18.5 (Wagner et al., 2005), the retina and developing lens vesicle

at E12 and the presumptive retinal ganglion layer between E15-postnatal day 1 (PD1) (Wagner et al., 2002a). Additionally it was found by studying the *lacZ* expression pattern driven by the human *Wt1* promoter in YAC transgenics that *Wt1* was expressed in the septum transversum (a structure from which the diaphragm develops), the coelomic mesenchyme and a domain between the developing digits of the limb (Moore et al., 1998). These sites of *Wt1* expression were verified as endogenous expression sites by whole mount RNA *in situ* hybridization.

Studies in the rat have also shown sites of *Wt1* expression by *in situ* mRNA hybridization to be in the developing nephrons and podocyte layer of the glomerulus of the kidney, the Sertoli cells of the testis, motor neurons in certain regions of the spinal cord and the area postrema of the brain (a structure within the fourth ventricle) (Sharma et al., 1992). They also report by Northern hybridization that *Wt1* was expressed in the kidney in the adult, in their figure however, a hybridization product in the lane corresponding to kidney RNA four weeks after birth (the latest stage tried) was not visible. Their statement that Call et al. (1990) did not detect *WT1* in human adult kidney was also unfounded, as it was in this paper that *WT1* cDNAs were isolated from these tissues; it is possible they came to this conclusion based on the Northern hybridization results which showed the presence of strong RNA transcripts in sporadic Wilms' tumour samples without mentioning the situation in normal adult tissue (Call et al., 1990).

Concentrating on expression in the murine urinary system, expression at E10 was confined to the mesonephric vesicle but not the mesonephric tubules (Armstrong et al., 1993). At E11 expression was also seen in the uninduced metanephric mesenchyme and later at E12 in the induced metanephric mesenchyme and subsequent mesenchymal condensations. At E13 there was little expression remaining in the regressing mesonephros. There was weak expression in the renal cortex whereas expression in the metanephros was high, with sites of expression being the metanephric vesicles, delimitating to the presumptive podocyte cells of the glomeruli where it persists in the adult. Within the genital system, expression at E10-E12 was strong over the genital ridge; by E13 when

the gonads are forming, expression was restricted to the sex cords. At E15 there was strong expression in the gonads, and the mesorchium membrane that supports the testis and mesonephros. In adult tissues expression was shown to be restricted to the Sertoli cells. Where human tissue was available, expression patterns correlated with those found in the mouse at corresponding developmental stages (Armstrong et al., 1993).

1.4.8 Roles of WT1

Mouse transgenic and knockout studies have allowed for the function of the differing isoforms of WT1 to be dissected (see section 1.4.15). Although there are a range of phenotypes which can be ascribed to each individual isoform by analysis of the mutant phenotypes, it is apparent that no one isoform or even all isoforms are entirely responsible for the Wilms' Tumour phenotype.

WT1 under different circumstances is a tumour suppressor gene, oncogene, transcriptional activator, transcriptional repressor and post-transcriptional regulator. These diverse roles, some of which are seemingly conflicting, makes the dissection of the action of this gene a complex process, particularly as there are so many different isoforms of this gene. At least one piece of evidence for each of these diverse roles is discussed below.

1.4.8.1 Tumour suppressor gene

As discussed in section 1.4.6, inherited mutations at the *WT1* locus followed by a subsequent inactivation of a second *WT1* allele lead to Wilms' tumours. This 'two-hit model' of tumourigenesis provides evidence that this gene acts as a tumour suppressor gene, that when inactivated leads to the cellular proliferation and subsequent tissue overgrowth characteristic of tumours. A characteristic of tumour suppressor genes are that they are able to suppress the activity of tumour cells. This hypothesis was tested and confirmed in a human Wilms' tumour cell line (RM1) where growth suppression was observed following transfection of CMV-driven murine *Wt1* expression constructs of the four 17aa and KTS isoforms (Haber et al., 1993).

1.4.8.2 Oncogene

The expression of *WT1* in tumours of tissues, such as the breast, that do not usually express this gene and the observation that no specific mutations of this gene have been found in these tumours suggest that *WT1* is acting as an oncogene under these circumstances. Evidence for this oncogenic role was found in tumour cell lines where *WT1* was knocked down by antisense oligonucleotides resulting in an inhibition of proliferation and increased apoptosis (reviewed in (Hohenstein and Hastie, 2006). It has also been noted that *WT1* was highly expressed in some Wilms' tumours (Pritchard-Jones et al., 1990) rather than being absent as would be expected for the inactivation of a tumour suppressor gene (reviewed in (Menke et al., 1998b). Additionally, *WT1* mutations occur in 14% of acute myeloid leukaemias, most exhibiting heterozygous mutations implying a dominant or dominant-negative mode of action (King-Underwood and Pritchard-Jones, 1998).

1.4.8.3 Transcription Factor

The role of WT1 as a transcription factor was initially tested by an array of luciferase reporter assays with many genes hypothesised as targets for either activation or repression by virtue of their promoter activity in cell lines cotransfected with WT1 isoform specific expression constructs. These genes were initially postulated as targets due to the similarity of their promoter sequences to a shared CG-rich consensus sequence binding site for WT1 KTS-, termed the Early Growth Response 1 (EGR1) consensus site, and *in vitro* molecular methods such as DNase I footprinting. The transcriptional activity of WT1 in these experiments suggested this gene was a promiscuous regulator with different binding activities for the separate isoforms. Many of these initial interactions however were found to be experimental artefacts (reviewed in (Menke et al., 1998a). Subsequent genetic and biochemical studies were carried out in more physiologically relevant settings to uncover the true extent of transcriptional activity by WT1. Many types of targets have been identified (Table 1.1) including growth factor genes, growth factor receptor genes, transcription factors, extracellular/secreted protein-coding genes and genes with other roles (reviewed in (Scharnhorst et al., 2001). WT1 has also been shown to

Table 1.1

WT1 Target Genes	
Gene Type	**Genes**
Growth Factors	IGF2 PDGF-A CSF-1 TGF-β Amphiregulin Inhibin-α Midkine MIS CTGF
Growth Factor Receptors	InsulinR IGF1R EGFR RAR-α
Transcription Factors	WT1 SF1 FREAC-4 EGR-1 c-Myb c-Myc N-Myc PAX2 DAX-1 SRY
Extracellular/Secreted protein-coding	Syndecan-1 Thrombospondin 1 NovH E-cadherin
Others	$G\alpha_{i-2}$ ODC MDR-1 HSP70 p21 BCL-2 Cu,Zn-SOD RbAp46 HTERT

Table 1.1. Shows some of the genes that can be regulated by WT1 and their gene type classification. Most are putative target genes whose interaction has not yet been verified *in vivo*. Table adapted from (Scharnhorst et al., 2001).

negatively regulate its own transcription (Rupprecht et al., 1994). Evidence for WT1 acting as a transcriptional activator and repressor are discussed below.

Transcriptional Activator

Splicing factor 1 (*Sf1*), a gene involved in development of the gonads, was found to be transcriptionally activated by Wt1; specifically the WT1 KTS- isoform was found to bind to the promoter of *Sf1* by DNase I footprinting, and cause its activation as demonstrated by cotransfection experiments with a WT1 KTS- expression vector and an *Sf1* promoter fragment linked to a luciferase reporter in a murine Sertoli cell line (TM4) (Wilhelm and Englert, 2002); this promoter binding could of course have been a transfection artefact. More physiologically relevant is that the expression of *Sf1* was demonstrated to be dependent on Wt1 by its absence in $Wt1^{-/-}$ embryos, as shown by RT-PCR analysis and in situ hybridization suggesting a mechanistic role for the cooperation of *Wt1* and *Sf1* in sexual differentiation (Wilhelm and Englert, 2002).

Transcriptional Repressor

Epidermal growth factor receptor (*EGFR*), a gene involved in epithelial cell development, was found to be transcriptionally repressed by *WT1* under the control of an inducible tetracycline expression system within two osteosarcoma cell lines (U2OS and Saos-2) (Englert et al., 1995). Prolonged expression of *WT1* caused an induction of apoptosis dependent upon the DNA binding domain of WT1 remaining intact - suggestive of the importance of the transcriptional properties of WT1 in the induction of programmed cell death. The expression levels of growth factor receptor genes involved in epithelial cell development and kidney development were monitored in these cell lines and upon inducible expression of *WT1*, *EGFR* was found to be repressed (Englert et al., 1995). This repression could well be what occurs *in vivo*: in the developing rat kidney *EGFR* is most highly expressed in early structures and ceases to be expressed in later structures in which *WT1* is highly expressed (Englert et al., 1995).

1.4.8.4 Post-transcriptional regulator

The role of WT1 as a post-transcriptional regulator is unclear, however there is some *in vitro* and *in vivo* evidence that WT1 can bind to RNA molecules and that WT1 is able interact with RNA processing factors.

WT1 was shown to bind to specific RNA sequences within exon 2 of the *Insulin-like growth factor 2* (*Igf2*) gene via the zinc fingers, with the WT1 KTS+ isoform having a greater binding affinity (Caricasole et al., 1996). Following RNaseA treatment, bound WT1 zinc finger GST fusion products and *in vitro* transcribed *Igf2* exon 2 was identified by electrophoretic mobility shift assays (EMSAs). To ascertain whether this interaction could occur *in vivo*, cell lines that were known to express WT1 were stained with a monoclonal WT1 antibody in order to identify the specific cellular localisation of this protein. Most protein appeared in the nucleus with the areas of highest localisation being those areas with the lowest DNA, and upon RNase treatment WT1 protein could no longer be detected in the characteristic nuclear pattern whereas DNase treatment had no effect. Together, these two pieces of evidence are suggestive of an association of WT1 with RNA. WT1 RNA has been shown to form a duplex with the RNA transcript of the *WT1-AS* gene (see section 1.4.12).

Using a yeast two-hybrid approach it was shown that WT1 binds directly to a splice factor U2AF65 (more strongly with the WT1 KTS+ isoform) and by splicing assays using nuclear extracts from *WT1* expressing M15 cells that this gene can associate with components of the splicing machinery - spliceosomes (Davies et al., 1998). Additionally, it has been shown that ~10-50% of endogenous murine Wt1 (Wt1 KTS+ and Wt1 KTS- isoforms) can shuttle between the nucleus and cytoplasm, indicative of a post-transcriptional role (Niksic et al., 2004).

1.4.9 WT1 interacting proteins

The different spatial and temporal roles of the WT1 isoforms suggest that these proteins are modulated by other factors. It is now known that there are many proteins that can interact with WT1 and exert an effect on its function (Table 1.2). The transcriptional activity of WT1 for example is known to be modulated by proteins that act as transcriptional coregulators. Some of these proteins have been shown to act upon WT1 by affecting the functionality of its transcriptional repression and activation domains. One such protein, BASP1, was identified in a screen specifically looking for WT1 interacting proteins. This protein is the

Table 1.2

WT1 Interacting Proteins	
Protein	**Function of Interaction**
WT1	Required as a prerequisite for transcriptional regulation by WT1
BASP1	Required for transcriptional suppression by WT1
p53	WT1 modulates the transcriptional regulatory properties of p53 and suppresses p53-mediated apoptosis
p63	Unknown
p73	Blocks DNA binding by WT1. Required for transcriptional regulation of both proteins.
SF1	Cooperate to increase transcription of Müllerian inhibiting substance
U2AF65	Putative splicing regulation function
WTAP	Putative splicing regulation function
PAR4	Blocks transcriptional activation and growth suppression by WT1
CIAO1	Blocks transcriptional activation by WT1
HSP70	Cooperate to inhibit cell proliferation
CBP	Required for transcriptional activation by WT1

Table 1.2. Some of the proteins that can interact with WT1 and their interaction function. Table adapted from (Scharnhorst et al., 2001).

subject of part of this thesis and further details on the interaction between WT1 and BASP1 can be found in section 1.5.7.

1.4.10 The *WT1* antisense transcript

In addition to the many isoforms of WT1, there is also an overlapping antisense transcript that produces an RNA that is not translated into protein. This transcript was initially discovered when looking for *WT1* cDNAs in order to clone the *WT1* gene. It was initially thought that these two divergently transcribed genes were non-overlapping and shared a bidirectional promoter. It is now known that they do overlap, at least in the human genome, have their own promoters and share a regulatory region. These progressive discoveries are discussed below.

The *WT1* antisense transcript was first identified and mapped to the human 11p13 locus by Bonnetta et al. (1990) who were looking for transcripts in Wilms' tumour tissue. This antisense transcript, initially named *WIT-1* (now known as *WT1-AS*), was found to be transcribed in a telomeric-to-centromeric direction, divergently and non-overlapping from its sense transcript, initially named *WIT-2* (now known as *WT1*), from a single CpG island – regions often associated with promoters (Huang et al., 1990). Although the 5' ends of the transcripts were not cloned, it was postulated that these two transcripts shared a bidirectional promoter due to the presence of two TATA boxes and a CCAAT box in the presumed intervening region. A combination of RNase protection assays (RPA) and anchored PCR predicted that *WT1-AS* was a transcript of between 2.1-2.7kb followed by a polyadenylation signal with three possible introns due to the identification of four cDNAs of varying lengths with identical 3' ends (Huang et al., 1990). Gessler and Bruns (1993) tried to map the positions of these exon/intron boundaries by RT-PCR using human foetal cDNA but only found a single splice site junction.

Although Huang et al. (1990) initially showed by RPA that the 5' most regions of *WT1* and *WT1-AS* do not overlap, it has since been shown by RPA and Northern hybridization of human kidney tissue and a human foetal kidney cDNA library, that some *WT1-AS* transcripts, 7-10kb in size, are overlapping with intron 1 and exon 1 of *WT1* and its upstream region, making it a true antisense

transcript (Campbell et al., 1994; Eccles et al., 1994). Campbell et al. (1994) also showed by RPA and primer extension followed by cloning and sequencing that *WT1-AS* has several transcriptional start sites and one alternatively spliced exon. A further spliced transcript was identified during a data mining screen of *WT1-AS* cDNA and EST foetal kidney libraries, and was subsequently confirmed by real time RT-PCR analysis to be present in spleen, foetal kidney, normal kidney and Wilms' tumour tissue, but not liver or brain; the other spliced transcripts were also present in these tissues (Dallosso et al., 2007). This further spliced transcript contained two splices, one short and one large, both downstream of the splice site previously identified by Gessler and Bruns (1993) (Figure 1.3B). It is evident from aberration of splicing in certain tumour tissues that splicing may play an important role in the functionality of the *WT1-AS* transcripts: six out of nine acute myeloid leukaemia samples and one out of four acute lymphocytic leukaemia samples that expressed *WT1-AS* have been shown to be aberrantly spliced (Dallosso et al., 2007).

In the mouse much less is known about the *Wt1-AS* transcript. RPA on mouse newborn kidney picked up a divergently transcribed antisense product overlapping at least part of *Wt1* exon 1 (Campbell et al., 1994). Gong et al. (2001) however, whilst initially mapping the murine *Wt1* locus did not find any divergently transcribed transcripts that were overlapping. RPA was carried out on a mouse metanephric cell line which expresses WT1 (M15 cells) and a range of organs from 5 day old mouse pups using probes specific for either the 5' end of *Wt1*, exon 1 of *Wt1* or the intron 1/exon 1 boundary of *Wt1*. Although divergently transcribed *Wt1* transcripts were detected in the M15 cell line, the kidney and spleen (not detectable in their published figure) with the probe specific for the 5' end of *Wt1*, there were no detectable protected fragments with the other probes (Gong et al., 2001). Three alternatively spliced mouse *Wt1-AS* expressed sequence tag (EST) clones that were identified from a newborn ovary-derived library and whose corresponding cDNAs were obtained from the IMAGE Consortium were used in RT-PCR reactions with primers specific for different regions of the *Wt1* locus. These transcripts showed overlap with *Wt1* exon 1 and part of intron 1, and a large initial splice of approximately 25kb, downstream of the major *Wt1* transcriptional start site (Dallosso et al., 2007).

In order to investigate the conserved nature of the *WT1-AS* sequence between species a 20kb 5' region of the human and mouse *WT1* gene was aligned. Displayed as a percentage identity plot (PIP) it was apparent that there were two regions of high homology between the two species present at the 3' termini of alternatively spliced human kidney *WT1-AS* transcripts. These regions were named Mouse Conserved Region I (MCRI) and Mouse Conserved Region II (MCRII). MCRI was located ~11-12kb upstream of *WT1* exon 1 and shared a 78% sequence identity over 506 nucleotides. MCRII was located ~4-5kb upstream of *WT1* exon 1 and shared 94% sequence identity over 289 nucleotides (CLIC Sargent Research Unit, University of Bristol, personal communication). In addition to its high sequence identity between human and mouse, MCRII was also shown to be conserved in the puffer fish *Fugu rubripes* (Colin Miles, personal communication) suggesting, because it is non-coding, that it has a regulatory role. Further unpublished work by the CLIC Sargent Research Unit showed by EMSA that the region as a DNA element can form multiple complexes in renal cells expressing high levels of WT1 protein, *WT1* RNA and *WT1-AS* RNA. Further analysis into this potential regulatory region was carried out within transient transgenic mouse embryos containing a transgene consisting of the MCRII region linked to a minimal promoter *lacZ* cassette (Colin Miles, personal communication). The expression profile of these embryos showed a partial overlap with the *lacZ* expression profile of the 470kb YAC transgenics (Moore et al., 1998) that contain the MCRII sequence, suggesting that MCRII itself is a regulator of *Wt1* expression (see section 1.4.15.2).

1.4.11 *WT1-AS* expression in kidney and Wilms tumours

A study into the relationship between Wilms' tumour pathology and the expression levels of *WT1* and *WT1-AS* by *in situ* hybridization revealed that both genes were co-ordinately expressed in the condensing blastema, S-shaped bodies and immature glomeruli of normal foetal kidney and the blastema and epithelia of certain Wilms' tumours; *WT1-AS* expression being a tenth that of *WT1* (Yeger et al., 1992). Wilms' tumours in which there were markedly higher levels of *WT1* and *WT1-AS* were typically homotypic tumours containing a high percentage of

blastema and epithelia compared to stroma, whereas Wilms' tumours in which there were markedly lower levels were typically heterotypic tumours containing a high percentage of muscle and squamous epithelium compared to blastema. It was demonstrated that this high expression was not just due to the presence of additional blastema, as one homotypic tumour had low expression levels of these genes and another with limited blastema tissue had normal expression levels (Yeger et al., 1992). Immunohistochemistry and further *in situ* hybridization of WT1 expression in normal foetal kidney (15-week) and mature neonatal human kidney samples showed that *WT1*, *WT1-AS* and WT1 co-localised within the condensing blastema, S-shaped bodies, glomeruli and podocytes. Expression of *WT1*, *WT1-AS* and WT1 was highest in the immature kidney and lowest in the mature kidney (Moorwood et al., 1998).

1.4.12 Regulation of *WT1-AS*

A candidate bidirectional promoter region of the overlapping *WT1* and *WT1-AS* genes was initially found by transient transfection assays of portions of genomic DNA into WT1 expressing tissue culture cells followed by DNase I protection assays. This region was narrowed down to a 1.3kb area of the human locus containing the transcriptional start sites of both genes (Campbell et al., 1994). Subsequently, a region with four times as much promoter activity was found in *WT1* intron 1 by luciferase reporter assays using constructs containing parts of intron 1 DNA sequences in the antisense orientation, transiently transfected into T5A1 cells (Malik et al., 1995). T5A1 cells were a stable 293 HEK cell line expressing elevated levels of *WT1*, specifically the full length WT1 KTS- isoform that had previously been used to demonstrate that WT1 autorepresses its own sense promoter (Malik et al., 1994). By this method and primer extension using foetal kidney RNA, the promoter region was narrowed down to 560 nucleotides of intron 1. The location of the transcriptional start site was found to also be within intron 1 confirming the position of the promoter (Figure 1.3B). Sequencing of this region showed no TATA or CCAAT boxes as is the case with the *WT1* promoter, but unlike the *WT1* promoter there were also no GC-rich areas. Two putative high affinity WT1 consensus binding sites were also

identified (Malik et al., 1995). A putative promoter region within the murine locus has yet to be identified.

In order to further characterise the properties of *WT1-AS* promoter transactivation by WT1, a series of cotransfections were carried out in 293 HEK cells with the *WT1-AS* promoter and murine Wt1 isoforms under the control of a constitutively active CMV promoter (Moorwood et al., 1999). As previously reported by Malik et al. (1995), the *WT1-AS* promoter was transcriptionally activated by full length WT1 KTS- isoform. Other WT1 isoforms (KTS+/17aa+, KTS+/17aa-, KTS-/17aa-) however had no significant effect, demonstrating the need for both the presence of the 17aa splice and the absence of the KTS splice for transactivation of the *WT1-AS* promoter (Moorwood et al., 1999). When mutations were introduced within the two putative high affinity WT1 consensus binding sites upstream of the transcriptional start site of the *WT1-AS* promoter region (Malik et al., 1995) activation of the promoter by full length WT1 KTS- was not significantly affected. This result was surprising in that it demonstrated that these binding sites were not necessary for transactivation of the *WT1-AS* promoter by WT1, suggestive that DNA binding is not the mechanism by which WT1 exerts its affect on the *WT1-AS* promoter (Moorwood et al., 1999).

Once identified, the activity of this human *WT1-AS* promoter region in relation to *WT1* was addressed. Luciferase reporter constructs containing the promoter region sequence were transiently transfected into T5A1 cells. The *WT1-AS* promoter region was shown to be transcriptionally activated by WT1 and negatively regulated by upstream (relative to the 5' end of the antisense promoter) *cis*-acting elements (Malik et al., 1995). This transcriptional activation by WT1 was surprising, as up to that time WT1 was known to transcriptionally repress promoters of many other genes.

It also appears that *WT1* and *WT1-AS* may reciprocally regulate each others expression. By stably transfecting the T5A1 cell line with constructs containing *WT1* exon 1 sequences in the antisense orientation linked to a luciferase reporter construct, *WT1-AS* was observed to affect the expression of *WT1*. In cell lines where *WT1-AS* levels were high it was shown that WT1 protein levels were

downregulated (Malik et al., 1995); however in cell lines where *WT1-AS* levels were still high, but lower than previously, perhaps representing more biologically relevant levels in relation to the sense:antisense transcript ratios, the WT1 protein levels in the majority of these cell lines were in fact elevated not decreased, with the lowest levels of *WT1-AS* corresponding with the highest WT1 protein levels and vice versa (Moorwood et al., 1998). This latter result is surprising as it suggests that *WT1* and *WT1-AS* are not acting in accordance with the dogma for sense and antisense genes, the dogma being that an increase in antisense expression leads to a decrease in the production of protein product from the sense transcript, presumably by overlapping RNA products forming a duplex that is targeted for degradation. Another experiment supporting this deviation from the dogma, and one set in a physiologically relevant environment has already been discussed in section 1.4.11: Moorwood et al. (1998) showed by *in situ* hybridization and immunohistochemistry of WT1 expression in foetal and mature neonatal human kidney samples that *WT1*, *WT1-AS* and WT1 co-localised within the same developmental structures and that all were at their highest and lowest levels at the same developmental stages. Although at present it is not known how this positive feedback is achieved, it has been postulated that *WT1-AS* may increase *WT1* transcription, stabilise the sense transcript or increase its rate of translation. Indeed, it has been shown that RNA:protein duplexes can be formed *in vitro*; *WT1-AS* was *in vitro* transcribed from a cDNA clone in the presence of [^{32}P]NTP and combined with resin bound poly-histidine tagged WT1 peptide fragments of the zinc finger region with and without the KTS isoform insertion. RNA binding affinities to the recombinant peptides were confirmed and measured by scintillation counter readings (Ye et al., 1996). In this experiment the binding affinities of both the WT1 KTS+ and WT1 KTS- recombinant peptides were found to be equivalent, which is in contrast to the isoform specific binding affinities found by Moorwood et al. (1999) (see above). *WT1-AS* has also been found by RT-PCR to be present in cytoplasm fractions of 7.92 cells derived from a paediatric renal rhabdoid tumour (mentioned in (Brightwell et al., 1997), demonstrating that *WT1-AS* is exported from the nucleus, an essential condition if RNA:protein duplex formation was to occur *in vivo*. Co-localisation studies carried out in cytoplasm fractions of these 7.92 cells has also found that RNA:RNA duplexes of *WT1* and *WT1-AS* can be formed.

RNase protection RT-PCR was carried out with primers specific for different regions of the *WT1* locus, and a product was obtained using exon 1 primers, indicating that an RNA duplex was present at this location due to its exemption to RNaseA activity, although it can't be known for sure if this duplex formation was physiological or a product of the extraction procedure (Dallosso et al., 2007).

1.4.13 Imprinting of *WT1-AS* and *AWT1*

The imprinting status of *WT1-AS* and *AWT1* (an isoform of WT1 with an alternative promoter and leader exon) has been assessed by observing the methylation status of the promoter region of *WT1-AS* and *AWT1* in Wilms' tumours with and without LOH at the *WT1* chromosome location 11p13 (Dallosso et al., 2004; Dallosso et al., 2007; Malik et al., 2000); LOH in Wilms' tumours being known to occur non-randomly with the preferential loss of the maternal allele (Schroeder et al., 1987). Bisulphite sequencing was used to assess regional methylation patterns by distinguishing between methylated and unmethylated CpG residues of bisulphite-modified DNA. It was shown that in normal human paediatric kidney, a domain of the *WT1-AS* and *AWT1* promoter region adjacent to a CpG island, termed the WT1 antisense regulatory region (ARR), was preferentially methylated on the maternal allele, whereas hypomethylation was seen in the majority of Wilms' tumours without LOH. RT-PCR was then carried out in order to observe allele specific expression. It was shown that in human foetal kidney *WT1-AS* and *AWT1* are biallelically expressed, whereas in normal human paediatric kidney *WT1-AS* and *AWT1* are monoallelically expressed from the paternal allele. In Wilms' tumours without LOH, *WT1-AS* and *AWT1* were biallelically expressed. It appears therefore that there is a relaxation of imprinting at the WT1 ARR differentially methylated region (DMR) in the disease state. This imprinting of *WT1-AS* and *AWT1* is in contrast to *WT1* which has been shown to be equivalently expressed from both the paternal and maternal allele in foetal kidney and Wilms' tumours (Little et al., 1992). These observations of genomic imprinting at the *WT1* ARR were backed up by methylation sensitive Southern hybridization of genomic DNA from BWS patients with and without uniparental disomy (UPD) of a

chromosomal region including 11p13 (Dallosso et al., 2004). Samples exhibiting a paternal UPD showed hypomethylation of the *WT1* ARR region compared to samples not exhibiting UPD. Methylation of the ARR was therefore strongly implicated in the epigenetic regulation of the *WT1-AS* and *AWT1* transcripts.

The *WT1-AS* promoter has also been shown to be methylated in ovarian clear cell adenocarcinoma. Both the *WT1-AS* and the *WT1* sense promoters were methylated in ~88% of clear cell and ~22% of serous cell adenocarcinoma tissue (Kaneuchi et al., 2005). By RT-PCR it was confirmed that there was no expression of *WT1-AS* and *WT1* in clear cell adenocarcinoma tissue where there was high methylation, but there was expression of these two genes in serous cell adenocarcinoma tissue where there was little methylation (Kaneuchi et al., 2005).

With regard to the imprinting status of mouse *Wt1-AS*, there is no evidence to suggest that the expression of this gene is monoallelic as is the situation in humans. In fact, the following experiment provides some limited evidence that it is biallelically expressed. The *small eye* or Sey^H mouse containing a mutation that included deletion of the *Wt1* locus induced by X-irradiation of oocytes (cited in (Glaser et al., 1990), was used to transmit the deletion from either the paternal or maternal allele in order to monitor the expression levels of *Wt1-AS* in the kidney from each allele. This was done by real-time PCR using primers specific for one of the *Wt1-AS* splices. Expression was in fact not significantly altered when expressed solely from either allele although the numbers of mice used in this experiment were small (Dallosso et al., 2007).

1.4.14 *WT1-AS cis*-regulatory regions

Further characterisation of the *WT1* ARR DMR by luciferase reporter assays have identified it to be a *cis*-acting transcriptional silencer element of the *WT1-AS* and *AWT1* promoters (Hancock et al., 2007) (Figure 1.3B). Luciferase reporter deletion constructs of the *WT1* ARR DMR region and surrounding sequence were transiently transfected into 293 HEK cells. A decrease in luciferase activity was seen with constructs containing sequences encompassing and immediately upstream (relative to the direction of *WT1* transcription) of the

putative *WT1* ARR. This region has also been shown to be a methylation-sensitive binding site for the imprinting regulator protein, CCCTC-binding factor (CTCF), a protein that is able to block communication between enhancers and promoters. This binding capacity was demonstrated *in vitro* by EMSAs, whereby CTCF protein bound precisely to the transcriptional silencer element, and verified *in vivo* by chromatin immunoprecipitation assays (ChIP) (Hancock et al., 2007).

A region upstream of *WT1* has been hypothesised to be a nuclear matrix attachment region (MAR) and a *cis*-regulatory region important for *Wt1* expression in the kidney. DNase I hypersensitive site mapping identified this region as one of 11 in the *WT1* locus spanning from the 5' upstream region to intron 1 (Scholz et al., 1997). These hypersensitive regions represent areas of open chromatin that are potential functional sites, for example nuclear MARs. To ascertain whether these sites were also nuclear MARs, nuclear matrix from 293 HEK cells (WT1 expressing) and HeLa cells (WT1 non-expressing) were probed with radiolabelled *WT1* genomic fragments. Two fragments were found to be matrix attached, one of which was a 1.4kb region 15kb (± 500bp) upstream of both the human and mouse *WT1* transcriptional start site, a small subregion of which when sequenced was consistent with a nuclear MAR. This 1.4kb region (upstream of the MCRI region) showed homology between the human and mouse genomes: a 20-30bp region showed a greater than 80% sequence identity, and a >60bp region showed a less than 40% sequence identity (Scholz et al., 1997). Further verification that this region is an important functional site was shown by the loss of *Wt1* expression in 293 HEK cells transfected with a YAC clone containing the murine *Wt1* locus with a deletion of this 1.4kb region. Transfection of the YAC clone without the deletion displayed comparable expression levels to the endogenous situation (Scholz et al., 1997).

1.4.15 *WT1* knock out mice

The first mouse model created to study the role of Wt1 carried a targeted mutation of the *Wt1* gene generated by homologous recombination in ES cells followed by blastocyst injection (Kreidberg et al., 1993). Specifically, the first

exon of *Wt1* and 0.5kb of the upstream promoter region were deleted generating what was presumed at the time to be a *Wt1* null knockout mouse. It is possible that this mouse model is not a complete null due to the recent identification of the WT1 isoform, AWT1, containing an alternative promoter and leader exon within intron 1 (Dallosso et al., 2004).

Heterozygous knockouts on a C57BL/6:129/Sv hybrid background did not display an overt phenotype whereas homozygous knockouts died between E13 and E15. In these *Wt1* homozygous null embryos kidneys and gonads failed to develop and developmental abnormalities were seen in the heart and lungs due to defects in the mesothelial linings of the cavities surrounding these organs and the diaphragm. In the developing kidney the ureteric bud had failed to form by E11.5, cells of the metanephric blastema underwent an accelerated rate of apoptosis and by E12 the metanephric blastema had degenerated therefore leading to an absence of mesenchymal cells, the Wolffian duct being the only structure remaining. In gonadal development of *Wt1* homozygous null embryos, mesonephric tubules formed normally but not as extensively as in wildtypes (later shown to be due to development of cranial but not caudal tubules (Sainio et al., 1997). However, the main defect which could be seen from E11 was a reduced thickening of the epithelium of the gonadal ridge and a subsequent halting of development. As an absence of kidneys is not fatal until after birth, lethality of homozygotes was suggested to be caused by the severe developmental abnormality observed in the heart. After E12 hearts were smaller than normal, particularly the ventricles which had thin muscle linings. In some embryos there was pericardial bleeding and most significantly edema of the heart in almost all embryos (Kreidberg et al., 1993).

When the *Wt1* deletion was outbred onto a mixed genetic background (CBA, SWR, 129/Ola, C57BL/6, BalB/c) the homozygous nulls had the same phenotype, however it was found that 40% of the heterozygous knockout mice died by 400 days after birth compared to 5% of wildtypes. Of those that had died, 80% of the heterozygous knockout mice had developed glomerulosclerosis, an overgrowth of fibrous tissue of the glomerulus, between 60-386 days after birth (none of the wildtypes had detectable signs of glomerulosclerosis) (Menke et al.,

2003). Other statistics relating to the heterozygous knockout mice declared that by 150 days after birth 100% of males and 50% of females had developed proteinuria (abnormally large amounts of protein in the urine); 63% of males and 38% of females had developed albuminuria (presence of albumin in the urine, a protein required to maintain correct osmosis pressure); and 88% of males and 50% of females had high plasma creatine levels (Menke et al., 2003). These diagnosis tests therefore provided a good indicator of incident kidney failure. Morphologically, mice with albuminuria showed thickening of the GBM and fusion of the podocytes (Menke et al., 2003). Expression of *Wt1* in the podocytes into adulthood therefore seems to be required to keep these cells in their normal functional state.

On a 129/Sv inbred background, although heterozygous knockout males were fertile, females were infertile, with halted mitotic development of fertilised embryos prior to implantation (Kreidberg et al., 1999). The presence of a modifier gene modulating *Wt1* was determined to be the cause, rather than a strain specific difference in *Wt1* or an additional mutation in a closely linked gene (Kreidberg et al., 1999). The presence of modifier genes was also suggested by a study of the *Wt1* deletion on a MF1:C57BL/6 F1 hybrid background where they found that *Wt1* homozygous nulls survived until birth (dying immediately after due to respiratory failure) whereas there were no live births on the C57BL/6 inbred background (Herzer et al., 1999). On this hybrid background, embryos were found to have defects in the optical and olfactory systems from E12-E12.5, with thinner than normal retinas, apoptosis of retinal ganglion cells, disruption of optic nerve fibre growth, thinner than normal sensory olfactory epithelium and fewer neuronal progenitor cells; eyes were also noted to be reduced in size at E18 (Wagner et al., 2002a; Wagner et al., 2005). Prolonged survival of *Wt1* nulls on an MF1 outbred background also allowed a study into organ development at E18.5 were it was found that embryos lacked adrenal glands and had severe spleen abnormalities due to enhanced apoptosis in primordial spleen cells (Herzer et al., 1999).

The *Wt1* null mouse is not the only model that has been generated to identify functions of Wt1. There have been a host of techniques used to separate the

distinct roles of different isoforms and to verify the importance of Wt1 by adding back its function. These models are summarized below and most are reviewed in Discenza and Pelletier (2004).

1.4.15.1 Sey^{Dey}

The first model of a *WT1* deletion was a spontaneous mutant called *Dickie's small-eye* or Sey^{Dey}, encompassing a chromosomal deletion of two genes, *Wt1* and *Pax6*, as observed in human WAGR syndrome patients (cited in (Glaser et al., 1990). Homozygotes die at E6 or earlier and 60% of heterozygotes die during gestation. The surviving heterozygotes have a reduced body size and develop aniridia in adulthood, but not other symptoms of WAGR syndrome such as nephroblastoma or nephrogenic rests (Glaser et al., 1990). As two genetic loci are affected in this mutant, it is not known which gene is responsible for these phenotypes. In humans, deletion of one copy of *PAX6* is sufficient for aniridia but deletion of one copy of *WT1* is not sufficient for the development of Wilms' tumours. Nevertheless, WAGR patients have a higher risk of developing Wilms' tumours as they tend to lose the other *WT1* allele (see section 1.4.6); no Wilms' tumours have ever been observed in this mouse model however. It has been postulated that organ pathogenesis in mice is less sensitive to gene dosage than in humans, or that the gene responsible for the urogenital defects seen in WAGR patients is not contained within the Sey^{Dey} deletion (Glaser et al., 1990). It is also possible that Wilms' tumours do not arise in mice due to the fewer number of cell divisions required in the development of murine kidneys compared to human kidneys due to their smaller size and shorter time-span of development, thereby providing fewer possibilities for the loss of a second *Wt1* allele (Glaser et al., 1990). More likely however is the role of *IGF2/H19*, two epigenetically regulated genes linked to the *WT1* locus within the human genome but not the mouse genome. Indeed, *IGF2* is known to be overexpressed in the majority of Wilms' tumours (Wang et al., 1996).

1.4.15.2 YAC transgenics

Transgenics containing yeast artificial chromosome (YAC) sequences coding for two different lengths of the human *WT1* gene (470kb and 280kb) and driving

lacZ expression via the *Wt1* promoter were created in order to observe detailed expression patterns of *WT1*, particularly in the developing heart (Moore et al., 1998). The 470kb YAC spanned the entire *WT1* locus and accurately mimicked the endogenous *Wt1* expression pattern, indicating that there was good *WT1* sequence conservation between mouse and human. The 280kb YAC was a truncated form of the 470kb YAC finishing at a position approximately 10kb downstream of the *WT1* termination site. Heterozygous YAC transgenics on a mixed genetic background were then crossed with *Wt1* knockout mice (Kreidberg et al., 1993) to determine if the severe phenotype of the *Wt1* null could be alleviated by complementation (Guo et al., 2002; Moore et al., 1999). Complementation partially occurred with the 280kb YAC: *Wt1* null embryos survived gestation with no observable heart or septum transversum defects between E13.5-E17.5 (Moore et al., 1999), and no retinal defects at E18 (Wagner et al., 2002a). Neonates died within 48 hours after birth however with the observed defects of the kidneys being the probable cause (Moore et al., 1999). The complementation this YAC sequence provided was assumed to be due to the presence of some, but not all *WT1* regulatory sequences. This shorter YAC was therefore thought to contain regulatory elements required for heart and eye development, but maybe lacking regulatory elements required for kidney development. Complementation was more effective with the 470kb YAC: *Wt1* null embryos showed a partial rescue of the kidney phenotype with embryos developing both kidneys (26%), a single kidney (14%) or no kidneys (60%). The neonates born without kidneys died within 48 hours. The neonates with either one or two kidneys suffered from congenital nephrotic syndrome and died within three weeks of birth (Guo et al., 2002). This longer YAC was therefore thought to contain regulatory elements involved in kidney development, although the partial rescue indicates that this element was either not the only element involved in the development of this organ or that it was not highly conserved between mouse and human. The more effective complementation with the 470kb YAC was assumed to be due to *WT1* 3' regulatory regions.

1.4.15.3 YAC ubiquitous expression

Extensive failed attempts to generate transgenics containing the 280kb YAC fragment with a CMV promoter driving the *WT1* translational start site suggested that ubiquitous expression of *WT1* was not compatible with viability (Menke et al., 1998a).

1.4.15.4 Exon 5 deletion

Surprisingly, targeted deletion of *Wt1* exon 5 which results in the absence of the 17aa mammal-specific isoform of Wt1 did not detectably affect development, fertility or lactation of either heterozygous or homozygous mice on a 129:C57BL/6 hybrid background (Natoli et al., 2002b). It has been postulated that the role of this exon 5 splice may manifest itself in disease states rather than during development (Natoli et al., 2002b).

1.4.15.5 Overexpression of mutant Wt1 in podocytes

Constructs containing a deletion resulting in a truncated Wt1 protein lacking zinc fingers three and four, thought to act in a dominant-negative manner (as found in Denys-Drash patients), were overexpressed in the podocyte cells of the kidney under the control of the podocyte specific *nephrin* promoter whilst the wildtype levels of Wt1 protein were kept intact (Natoli et al., 2002a). These transgenes either contained or omitted exon 5. Transgenic mice that contained exon 5 within the transgene were born in the expected Mendelian frequency and appeared normal, showing no detectable developmental defects. Transgenic mice that omitted exon 5 within the transgene were rarely born; at E18 the embryos were present in the correct ratios but most exhibited abnormal glomeruli development (Natoli et al., 2002a). As *Wt1* DNA elements and Wt1 protein levels were intact in these transgenic mice, and the absence of exon 5 alone within the transgene did not result in a mutant phenotype, it was proposed that the neonatal deaths were due to a dominant-negative function of the mutant Wt1 peptide and that Wt1 isoforms lacking exon 5 are either exclusively involved or involved to a greater extent in glomerular development than isoforms containing exon 5 (Natoli et al., 2002a).

1.4.15.6 Denys-Drash model (1)

A mouse model mutation of WT1 (tmT396) with the aim of mimicking the DDS phenotype was generated by a targeted point mutation insertion into exon 9 of *Wt1*. This insertion created a truncated Wt1 protein product lacking functionality of the third and fourth zinc-finger and the KTS insert (Patek et al., 1999). In the original study 129/Ola-derived E14 and CGR8 targeted ES cells were injected into F2(C57BL/6JLac x CBA/CaLac) host blastocysts and chimeras were crossed with the inbred 129/Ola strain. Although only a single heterozygous male was derived from this line (born from a female chimera), the mouse was sterile and exhibited other features of DDS including diffuse mesangial sclerosis. Upon examination of the urogenital system of chimeras following autopsy, abnormalities were found in the kidneys, including glomerulosclerosis and in one case a Wilms' tumour, and immature Sertoli cells were found in the testis, the latter finding perhaps providing an explanation for the absence of progeny containing the targeted allele transmitted from male chimeras (Patek et al., 1999). In a later study 129/Ola-derived HM-1 ES cells were injected into C57BL/6JLac host blastocysts and chimeras were crossed with the inbred 129Ola or C57BL/6 strains (Patek et al., 2007). Male germline heterozygotes were then bred to three strain backgrounds: inbred 129/Ola, crossbred C57BL/6 x 129/Ola and outbred MF1. Apart from inbred females, the heterozygous mice were fertile, but all developed kidney disease. Homozygous mice died around E12.5 with the appearance of heart and diaphragm defects, hypoplastic livers, small adrenal primordia and an absence of metanephroi and gonads (Patek et al., 2007).

1.4.15.7 Denys-Drash model (2)

A common missense mutation found in DDS patients (R394W) resulting in a nucleotide and amino acid change in Wt1 was introduced into ES cells followed by blastocyst injection (Gao et al., 2004). $Wt1^{R394W/R394W}$ homozygous mutant mice were embryonic lethal whereas $Wt1^{R394W/+}$ heterozygous mutant mice were developmentally normal on an F1 hybrid C57BL/6:129 background and an MF1 outbred background. On an outbred background however, 100% of male mice and 42% of female mice had developed proteinuria and mesangial sclerosis by four months of age. The differences between the phenotypes of the mutation on

different genetic backgrounds points to the presence of a *Wt1* modifier gene. Upon investigation of Wt1 expression levels between the two strains it was found that there was 50% less Wt1 expression on the outbred background (Gao et al., 2004). This observation that the mutation is more severe phenotype on an MF1 outbred background is interesting as it is in contrast to the effect of a less severe phenotype of the $Wt1^{+/-}$ knockout mice when on this same outbred background (Herzer et al., 1999); perhaps this is due to the proposed dominant-negative function of the mutant DDS Wt1 peptide (see section 1.4.15.5).

1.4.15.8 Frasier Model (1)

The molecular defect in FS is a mutation proposed to cause a shift in the ratios of WT1 KTS+ and WT1 KTS- isoforms. FS patients show focal segmental glomerular sclerosis and male-to-female sex reversal. In order to generate a model of FS, mice lacking the Wt1 KTS+ or Wt1 KTS- isoforms were generated by targeted mutations within ES cells followed by blastocyst injection to create breeding chimeras (Hammes et al., 2001). Heterozygotes with only one functional copy of the Wt1 KTS- isoform developed normally and were fertile, whereas 70% of heterozygotes with only one functional copy of the Wt1 KTS+ isoform died between 2-3 months of age due to the development of albuminuria and glomerulosclerosis. They did not appear to have genital malformations however, and were fertile. This phenotype was shown to be absent on a different strain background, at least up until 6 months of age, suggesting the influence of modifier genes. Both Wt1 KTS- mutant homozygotes and Wt1 KTS+ mutant homozygotes died within 24 hours of birth due to kidney agenesis (Hammes et al., 2001), demonstrating the importance of both alternatively spliced isoforms for normal kidney development. Homozygotes for the Wt1 KTS- isoform mutation also exhibited gonads of a drastically reduced size with abnormal development of the genital ducts and a lack of differentiated gonadal tissue. Homozygotes for the Wt1 KTS+ isoform mutation showed a male-to-female sex reversal with the gonads of XY females showing reduced and disorganised germ cells (Hammes et al., 2001). A later study using these mice found that as with *Wt1* null mice, the sensory olfactory epithelium was thinner than normal in Wt1 KTS+ deficient mice from E12.5 (Wagner et al., 2005).

1.4.15.9 Overexpression of Wt1 KTS- isoform

Transgenic mice expressing an excess of the Wt1 KTS- isoform were generated by the integration of a transgene containing this cDNA isoform under the control of regulatory elements known to drive expression in the kidney and other tissues of the developing embryo (Lahiri et al., 2007). Three independent transgenic lines were generated, all expressing additional low levels of the Wt1 KTS- isoform and all appearing to be developmentally normal and fertile on a C57BL/6:CBA hybrid background. To address whether overexpression of the Wt1 KTS- isoform alone was sufficient for normal development, these transgenic mice were crossed with heterozygous *Wt1* null mice (Kreidberg et al., 1993) followed by intercrossing of the resulting double mutants. The presence of the Wt1 KTS- isoform was not sufficient to rescue the developmental abnormalities and subsequent prenatal lethality of *Wt1* null mice. In fact, an excess of Wt1 KTS- predisposed heterozygous *Wt1* null mice to kidney defects, with a high proportion of the male double mutants developing proteinuria from four months of age with corresponding cystic kidneys, accompanied occasionally by liver tumours and cystic testis resulting in sterility (Lahiri et al., 2007). This phenotype, in cells that normally express *Wt1*, has been attributed to a change in the KTS- : KTS+ isoform ratios.

1.4.15.10 Knockout of Wt1 alternative translational start isoform

In order to ascertain the role of the *WT1* isoform with an alternative translational start site, a mouse model was generated to knock-out this mammal specific isoform. ES cells were targeted with a translational stop signal 12 amino acid downstream of the upstream *Wt1* CUG alternative translational start site. A selectable marker was also inserted into a region of *Wt1* intron 1 not highly conserved between human and mouse and was subsequently removed by a cre-mediated *loxP* recombination leaving a single remaining *loxP* site (Miles et al., 2003). Mice containing this mutation were then generated by blastocyst injection to generate breeding chimeras. Both heterozygous and homozygous absence of these alternative start site isoforms on a C57BL/6 genetic background resulted in

no apparent developmental abnormalities and mice were fertile (Miles et al., 2003).

1.4.15.11 Compound mutants: *Wt1* knockout crossed with *p53* knockout

As Wt1 and p53 can alter each others transcriptional activities (Maheswaran et al., 1993) it was decided to cross *Wt1* heterozygous and homozygous null mice (Kreidberg et al., 1993) with *p53* homozygous null mice (cited in (Menke et al., 2002) in order to investigate their genetic interactions (Menke et al., 2002). On a mixed genetic background both $p53^{-/-}$ mice and $Wt1^{+/-}/p53^{-/-}$ mice developed lymphomas (90%) and sarcomas (15%), although $Wt1^{+/-}/p53^{-/-}$ mice showed an earlier onset. $Wt1^{+/-}/p53^{-/-}$ mice also developed epithelial neoplasm of the kidney (20%) and glomerulosclerosis (3%). $Wt1^{-/-}/p53^{-/-}$ embryos did not survive past E11.5, dying several days earlier than $Wt1^{-/-}$ embryos; both knockouts presumably dying from the observed pericardial bleeding (Menke et al., 2002). The absence of p53 does not therefore change the phenotype of mice that have an absence of Wt1.

1.4.15.12 Compound mutants: *Wt1* knockout crossed with Pax2 protein truncation model

As Wt1 and Pax2 were thought to regulate each others transcriptional activities during renal development, $Wt1^{+/-}$ heterozygous knockouts were crossed with $Pax2^{1Neu/+}$ heterozygous mutant mice in order to examine the interaction of these two genes (Discenza et al., 2003). $Pax2^{1Neu}$ mice have a targeted insertion leading to a frameshift mutation resulting in an inactive truncated Pax2 protein. $Pax2^{1Neu/1Neu}$ homozygous mice die immediately after birth due to an absence of kidneys, whereas $Pax2^{1Neu/+}$ heterozygous mice exhibit abnormalities of the kidney, eye, ear and brain; kidneys being 20% smaller than the kidneys of wildtype mice (Favor et al., 1996). The kidneys of $Wt1^{+/-}/\ Pax2^{1Neu/+}$ mice were 50% smaller than the kidneys of wildtype mice, and histological examination revealed a reduction in the size of the renal medulla, reduced development of the renal pelvis and fewer renal calyces, collecting ducts and glomeruli compared to $Pax2^{1Neu/+}$ mice (Discenza et al., 2003). It appears from the changes noted in the

Pax2 mutant phenotype when *Wt1* is also absent, that *Wt1* was behaving as a modifier gene.

1.4.15.13 *Wt1* conditional mutant

As *Wt1* null mutants die before or shortly after birth depending on the strain background, conditional mutants are needed to analyse Wt1 deficiencies in specific tissues. In order to ascertain the role of Wt1 in testis development, a conditional mutant was generated that only produces Wt1 peptides lacking zinc fingers 2 and 3 specifically in the Sertoli cells of the testis (Gao et al., 2006). This conditional mutation did not act in a dominant-negative manner but was a null mutation. This was achieved by flanking exons 8 and 9 of *Wt1* with *loxP* sites to produce $Wt1^{+/flox}$ mice. $Wt1^{+/flox}$ mice were then crossed with $Wt^{+/-}$ knockout mice (Kreidberg et al., 1993) to produce $Wt1^{-/flox}$ mice and finally, the $Wt1^{-/flox}$ mice were mated with the *AMH-Cre* transgene mice that specifically express *Cre* at E14.5 in the Sertoli cells of male mice (Lecureuil et al., 2002 cited in Gao et al., 2006). The testes of $Wt1^{-/flox}$;*AMH-Cre* mice were only 10% the size of those in wildtype mice and lacked seminiferous tubules resulting in a loss of both Sertoli cells and germ cells. The entire male reproductive tract was underdeveloped and in addition featured a uterus and structures resembling an oviduct. Considering these phenotypic observations, it was postulated that Wt1 is required for the maintenance of seminiferous tubules and Sertoli cells in the testis and for the regression of the Müllerian duct system rather than differentiation of the Wolffian duct system (Gao et al., 2006).

1.4.16 Other genes involved in Wilms' tumour

Although much has been discovered about the genetics of Wilms' tumour in relation to the WT1 gene, only 10-15% of all Wilms' tumours are considered to be due to an aberration in the coding region of this transcriptional factor (reviewed in (Hastie, 1994). Therefore, it is important to continue to investigate other possible genes involved in this disease and the regulation of WT1.

1.4.17 Other Wilms' tumour loci

Chromosomal rearrangements and allele loss in Wilms' tumours have pointed to several loci involved in the occurrence of this disease. 11p13 is the locus in which *WT1* can be found and 11p15, 16q, 1p, 7p, 11q, 22q and the X chromosome are other mutated loci associated with the disease (reviewed in (Hastie, 1994; Rivera et al., 2007).

The 11p15 locus was also found to involved in BWS patients and the gene responsible has been hypothesised to be imprinted due to specific maternal LOH of the 11p region in Wilms' tumours and the observation that 20% of BWS patients exhibit paternal UPD. *IGFII* for example is overexpressed in Wilms' tumours due to LOI (loss of imprinting) and overexpressed in BWS due to LOH. Inactivating mutations in $p57^{KIP2}$ have also associated with BWS (reviewed in (Ward, 1997). Good evidence for the involvement of both of these genes were mouse models displaying the BWS phenotype containing a knockout of $p57^{Kip2}$ and or an overexpression of *IfgII* (Eggenschwiler et al., 1997).

LOH also occurs in tumours at 16q with a frequency of 20%, and again, it is not known which gene is responsible for Wilms' tumours on this chromosome (reviewed in (Hastie, 1994). LOH in tumours at the X chromosome occurs with a frequency of approximately 33% and has been attributed to a recently identified gene named *WTX*. Due to the gene being located on the X chromosome, it only takes monoallelic loss to completely inactivate its function. Wilms' tumours in which there is a *WTX* mutation lack mutations in *WT1*. (Rivera et al., 2007). A mutation at 3p22 that activates the β-catenin gene *CTNNB1* has also been associated with Wilms' tumours. *CTNNB1* mutations are often found to coincide with *WT1* mutations in Wilms' tumours.

1.5 The BASP1 gene

1.5.1 Introduction to BASP1

The protein BASP1 was independently discovered by several groups working on different vertebrate species and has consequently been named more than once.

These proteins were gradually identified as homologues by peptide sequencing analysis. BASP1 was identified in the chick by Widmer and Caroni (1990) and was referred to as CAP23 in their communication and subsequent work. BASP1 was identified in the rat by Maekawa et al. (1993) and was referred to as NAP22, and also by Mosevitsky et al. (1994) who named the protein BASP1. A full historical account of the protein's discovery can be found in section 1.5.3.

The alias BASP1 has been chosen for use throughout this review (save for sections specific to the independent discovery of the protein in different species) as it is the protein name appearing most frequently in the commonly used sequence databases: GenBank at the National Centre for Biotechnology Information (NCBI), European Molecular Biology Laboratory (EMBL) and DNA Data Bank of Japan (DDBJ).

1.5.2 Properties of the family of brain acid-soluble proteins including BASP1

Brain acid-soluble proteins (BASPs) are a group of proteins found to be highly abundant in the cytoplasmic fractions of rat brain homogenates, most abundantly in the cortex and hippocampus (Mosevitsky et al., 1994).

This protein group was isolated during a screen for novel brain-specific proteins involved in biochemical processes of brain function. Four proteins were identified by their differing migration patterns during acetic acid-urea polyacrylamide gel electrophoresis (PAGE) of rat cortex samples (Mosevitsky et al., 1994). Each of these proteins shared properties of being very acidic and therefore hydrophilic with mean isoelectric points (p*I*) between 4.4-4.6. They also had low mobility during sodium dodecyl sulphate-PAGE (SDS-PAGE), had similar relative molecular masses (20-24kDa) determined by electrophoresis, were soluble in 5% perchloric acid and were all seemingly phosphorylated as indicated by staining of the protein bands for phosphorus. Many of these physiochemical properties are also shared with the well characterised neuronal growth associated protein 43 (GAP43) (Coggins and Zwiers, 1991) and myristoylated alanine-rich C kinase substrate (MARCKS) (Blackshear, 1993). Immunological tests using anti-GAP43 monoclonal antibodies have identified

BASP1 and BASP3 as novel proteins whereas BASP2-1 and BASP2-2 were identified as two forms of GAP43 (Mosevitsky et al., 1994). Not all of the BASPs therefore are evolutionary conserved proteins.

The novel protein BASP1 showed decreased mobility in non-ionic detergent (Triton X-100) highlighting the possibility of this protein containing a hydrophobic region, which was later shown to be the case due to the presence of an N-terminal myristic acid residue (Mosevitsky et al., 1996; Mosevitsky et al., 1997; Takasaki et al., 1999). These myristoylated proteins can localize and anchor to the lipid membranes. Specifically, BASP1 localizes in a cholesterol-dependent manner to membrane domains within the brain, called rafts, which form in the exoplasmic leaflet of the lipid bilayer (Iino et al., 2004; Maekawa et al., 1999; Terashita et al., 2002). Membrane rafts can selectively control the import and exclusion of proteins across the cell membrane (Salaun et al., 2004; Simons and Toomre, 2000), and it has therefore been postulated that BASP1 may be involved in this dynamic membrane control but no evidence has yet been put forward to support this hypothesis. Data suggesting an involvement of BASP1 together with GAP43 and MARCKS in actin regulation and neurite outgrowth at plasmalemmal microdomains by sequestering and modulating phosphatidylinositol-4,5-bisphosphate ($PI(4,5)P_2$), a signalling phospholipid of the plasma membrane, may be the first step in understanding the involvement of BASP1 in raft dynamics (Laux et al., 2000).

BASP1 was also shown to be a calcium-dependent calmodulin-binding signalling protein and a prominent substrate of protein kinase C (PKC) (Maekawa et al., 1993; Maekawa et al., 1994). This interaction with calmodulin was a direct consequence of fatty acid modification (myristoylation), with calmodulin binding to a domain consisting of part of the myristoyl group together with nine amino-acids of the N-terminus (myr-GGKLSKKKK) (Takasaki et al., 1999). Phosphorylation of the Ser^6 residue of BASP1 has been shown to inhibit this binding of calmodulin when phosphorylated with PKC *in vitro* and conversely, binding of calmodulin inhibits phosphorylation (Maekawa et al., 1994).

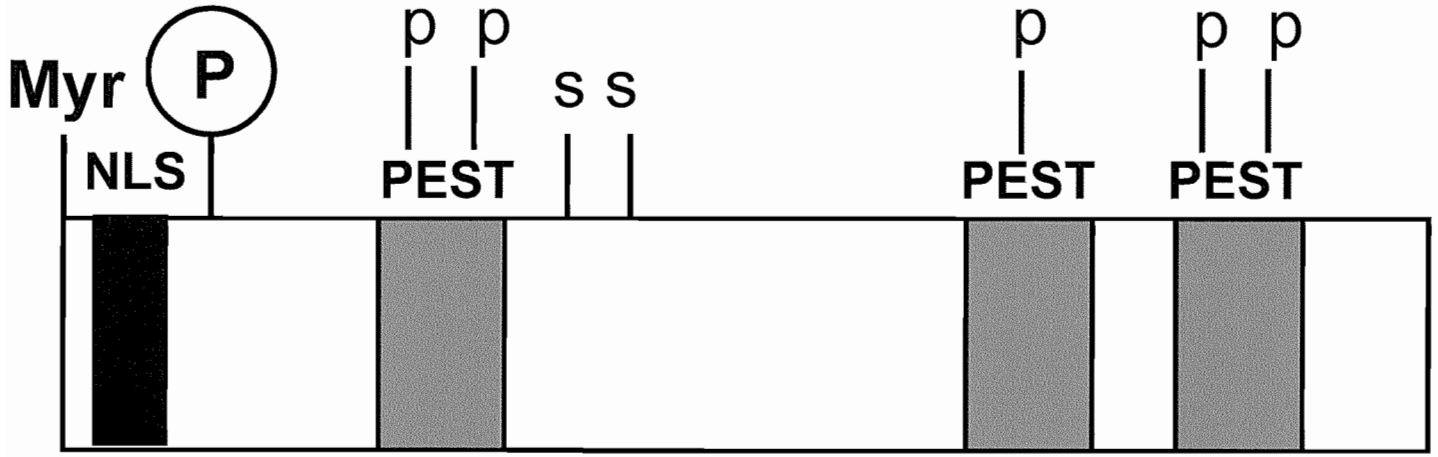
Not all BASP1 proteins are myristoylated however, as some protein isoforms, most likely to have arisen due to alternative splicing, are devoid of the N-terminal portion of the protein (Carpenter et al., 2004). Additionally, BASP1 has been shown to contain other motifs of unknown functional significance. These motifs consist of an N-terminal nuclear localization signal (NLS) (Carpenter et al., 2004), several 'high score' PEST sequences within the C-terminal region characteristic of short lived high-turnover proteins, and several potential PKC and casein kinase II (CKII) phosphorylation sites nested within the PEST sequences (Mosevitsky et al., 1997) and two potential sumo-modification sites (Stefan Roberts, personal communication) (Figure 1.4).

1.5.3 Identification of BASP1

BASP1 was initially identified in the chick by Widmer and Caroni (1990) whilst searching for proteins with similar physical properties and subcellular localisation as the brain-specific calmodulin-binding protein GAP43. Protein was extracted from embryonic day 13 (E13) chick brain and subjected to SDS-PAGE. Any proteins exhibiting abnormally low mobility during SDS-PAGE (a property of GAP43) were recovered by gel purification. One such protein, with an approximate molecular mass of 23kDa was BASP1 (termed cortical cytoskeleton-associated protein 23 or CAP23 by Widmer and Caroni (1990)), a novel cytosolic protein that like GAP43, was shown to be an acidic PKC substrate. The primary structure of this novel protein was deduced after the isolation and subsequent nucleotide sequencing of cDNA clones coding for chick CAP23 (Widmer and Caroni, 1990). This amino acid sequence was then compared against the known proteins in the EMBL and GenBank databases and found to indeed be novel, the highest homology being with GAP43, both proteins showing unusual highly charged, hydrophilic amino acid compositions (Widmer and Caroni, 1990; Wiederkehr et al., 1997).

BASP1 in the rat was identified by Maekawa et al. (1993), a group interested in molecules present in the growth cone fractions of embryonic rat brains. An acidic, hydrophilic protein with an approximate molecular mass of 22kDa was detected during SDS-PAGE and subsequently purified by column fractionation

Figure 1.4. Schematic representation of the domains of BASP1. The schematic shows the approximate localisation of the three 'high-score' predicted PEST sequences, the myristoylation signal (Myr), phosphorylation site (P), potential protein kinase C and casein kinase II phosphorylation sites (p), potential sumo-modification sites (s) and the nuclear localisation signal (NLS) close to the N-terminus.


Myr
P
NLS
p p
PEST
s s
p
PEST
p p
PEST

and termed neuronal axonal membrane protein 22 or NAP22. This protein of unknown function showed abnormal migration behaviour during SDS-PAGE, similar to that of GAP43 and MARCKS and like these proteins was shown to be able to bind calmodulin, this binding activity being in a calcium-dependent manner. cDNA cloning of this gene identified a possible myristoylation site at its N-terminus (Maekawa et al., 1993), and sequence comparisons showed that the protein has an almost identical N-terminal 50 amino acid sequence to that of the previously identified chicken CAP23 protein (Widmer and Caroni, 1990). The C-terminus of these two proteins was less similar however, leading to a sequence similarity over the entire peptide length of 65% between rat NAP22 and chicken CAP23 (Maekawa et al., 1993). This sequence evidence along with shared biochemical properties between the two proteins, lead to the authors to postulate that rat NAP22 was in fact either a rat homologue of chick CAP23, or a member of the same gene family. Several years later, Caroni et al. (1997) found sequencing errors in their previously reported chicken CAP23 sequence (Widmer and Caroni, 1990), specifically two translational errors caused by frame shifts mistakenly introduced at the 3'-end sequence of the cDNA. Park et al. (1998) who set out to characterize the cDNAs encoding the bovine and human NAP22 homologues also found the errors in the chicken CAP23 sequence and showed that by shifting the reading frame the C-terminal region had a much higher amino acid identity to that of the mammalian NAP22 protein sequences. Comparisons of homologies after this frame-shift correction, along with expression pattern comparisons between species after analysis of bovine NAP22 within tissues suggested that the four proteins were likely to have similar cellular functions. As with NAP22 and CAP23, the bovine and human NAP22 proteins showed the same almost identical N-terminal 50 amino acid sequence, suggesting an important role for this region in protein function (Park et al., 1998).

After the identification of the family of BASP proteins in rat brain homogenates by Mosevitsky et al. (1994), the same group went on to determine the amino acid sequence of rat, bovine and human BASP1 by the Edman N-terminal sequencing method (Mosevitsky et al., 1996). After a partial sequence of rat BASP1 was obtained, it was found by database protein sequence comparisons that regions of the rat BASP1 protein were very similar to regions of the previously sequenced

NAP22 rat protein (Maekawa et al., 1993). Rat BASP1 and NAP22 were therefore considered to be the same protein. Sequence comparisons of the bovine and human BASP1 amino acid sequences with the protein databases revealed that they were in fact homologues of both chicken CAP23 and rat NAP22.

Although peptide sequence homology between species was high for some regions of the BASP1 proteins, surprisingly polyclonal antibodies were species-specific indicating that the antigenic determinants must differ between species and may lie outside the homologous regions; BASP1 was therefore considered to have high evolutionary variability (Mosevitsky et al., 1996; Mosevitsky et al., 1994). Percentage identities of amino acids over the entire peptide sequences between human BASP1 and bovine BASP1, rat BASP1 (NAP22) and chicken BASP1 (CAP23) were found to be approximately 70%, 65%, and 56% respectively (Park et al., 1998) (further protein and sequence alignments are presented in sections 3.4.1 and 3.4.2).

Proteins immunologically related to BASP1 by recognition with anti-BASP1 antibodies, found to be comprised of N-terminal myristoylated fragments of this protein were identified by Zakharov et al. (2003) and termed BASP1 immunologically related proteins (BIRPs). These proteins were found to be present in all tissues in which BASP1 was located and were of a similar size in each species examined, and hence deemed to be of physiological significance (Zakharov et al., 2003). The basis on which these conclusions were made seem premature however, and data to qualify some of the claims was omitted.

1.5.4 BASP1 protein and RNA transcript size

Due to the independent discovery of this protein by several groups, the protein size was also determined on numerous occasions, with varying conclusions. When the proteins from whole brain rat homogenates (P3-P8) were separated by SDS-PAGE, the BASP1 protein was estimated to be 58kDa when separated in 7.5% acrylamide and 45kDa when separated in 12.5% acrylamide (Maekawa et al., 1993); much larger than the size of the protein determined from the predicted amino acid sequence, which was only 22kDa when deduced from the rat

nucleotide sequence (Maekawa et al., 1993). Mosevitsky et al. (1994) discussed the apparent change in molecular mass of BASP1 depending upon the concentration of acrylamide gel used to separate the proteins. When BASP1 isolated from rat brain homogenates was separated by SDS-PAGE in acrylamide concentrations ranging from 6%-22%, the relative molecular masses were determined to be 77-32kDa, with the protein migrating faster in a higher concentrations of acrylamide, most likely because of weak associations with SDS due to the high hydrophilicity of this protein (Mosevitsky et al., 1994; Widmer and Caroni, 1990). An independent estimation of the rat BASP1 protein size was calculated to be 21kDa, by comparing the relative mobilities of BASP1 against standard proteins such as egg albumin (Mosevitsky et al., 1994). Widmer and Caroni (1990) whilst isolating BASP1 from chick brain homogenates also noticed differing migration patterns during SDS-PAGE with varying concentrations of acrylamide, they estimated the size of the protein to be 23kDa, but how they came to this conclusion is unclear. When the human and bovine BASP1 proteins were sequenced the molecular mass calculated from their amino acid composition was 23kDa, 210 daltons lower than that measured by spectrometry (Mosevitsky et al., 1996).

In vitro studies using polyclonal antibodies raised against BASP1 were used to immunoblot several cell lines separated by 8-18% gradient SDS-PAGE in order to determine the size of the BASP1 protein within these cultured cells (Carpenter et al., 2004). These Western blots showed two BASP1 species migrating at 40kDa and 52kDa in HeLa cell nuclear extract, monkey kidney Cos-1 cells and HEK 293 cells. Human paediatric kidney tumour G-401 cells contained mainly the 40kDa form of BASP1. Mouse kidney M15 cells contained both forms of the protein (although in variable ratios due to heterogeneity of the cell extracts) plus a larger immunoreactive protein. Finally, the human erythroleukemia K562 cell line did not contain any BASP1 protein. The larger immunoreactive protein was suggested to be myristoylated due to the observation that the addition of an inhibitor of protein myristoylation (Triacsin C) to M15 cells shows a significant decrease in this slow-migrating protein product. The 40kDa and 52kDa forms of BASP1 on the other hand were considered to be non-myristoylated, as a significant increase in the levels of these proteins were seen when N-terminally

haemagglutinin (HA)-tagged BASP1-coding sequences under the control of CMV were transfected into M15 cells, this tagging preventing N-terminal myristoylation (Carpenter et al., 2004).

Other than nucleotide sequence data, evidence to ascertain the length of BASP1 transcripts is limited. Widmer and Caroni (1990) showed by northern analysis of rat head and brain samples, the presence of an mRNA species they estimated to be approximately 1.5kb in size and an additional minor hybridizing species whose estimated size was not mentioned but was smaller than the main RNA species. In a separate study, northern analysis of bovine tissue samples claims to show a transcript of 1.7kb, although the blot presented was of poor quality (Park et al., 1998).

1.5.5 BASP1 expression and localization

Patterns of BASP1 protein expression within the chick were detected by Western blotting of tissue homogenates or immunofluorescence of tissue sections (Widmer and Caroni, 1990). The distribution of BASP1 mRNA was measured by Northern blotting and was found to have an extremely similar distribution to that of the BASP1 protein. At E2 in the chick embryo, BASP1 was detectable in all tissues, but by E4 differences in expression patterns became apparent. In the developing liver, heart, intestines and limbs the levels of BASP1 remained high at E4, whereas there was almost no detectable BASP1 protein in other regions of the embryo. At further developmental stages, BASP1 became restricted to a narrower range of cells within the organs in which it was expressed, with levels of the protein no longer detectable in the liver at E12 or in the skeletal muscles at E18. At birth significant levels of the protein were only detectable in the intestines and central nervous system (CNS) (Widmer and Caroni, 1990). After birth, there was no detectable change in expression within the nervous system until after postnatal day 8 (PD8), when downregulation of BASP1 occurred, with no detectable BASP1 after P14 (Caroni et al., 1997). This data suggests a role for BASP1 during early organ development and an additional role within the nervous system.

Immunoblotting using a specific monoclonal antibody against this protein has shown BASP1 expression in rat tissue extracts isolated from the growth-cone fractions of the brain (Maekawa et al., 1993). Later studies have shown specific expression in the nerve terminals of the CNS from birth through to adulthood, specifically in the synaptic vesicles, presynaptic and postsynaptic membranes, and the outer mitochondrial membranes of the brain and spinal cord, with quantities increasing and finally accumulating four weeks after birth (Iino et al., 1999; Iino and Maekawa, 1999; Maekawa et al., 1993; Yamamoto et al., 1997). BASP1 is not only found to be expressed in the brain however, but is in fact also expressed more widely in other tissues such as the thymus, spleen, liver, kidney, testis, lungs, heart, tongue and CNS, as shown by SDS-PAGE of rat protein samples; Western blotting; peptide maps (Mosevitsky et al., 1996) (Novitskaya et al., 1994); and/or immunohistochemistry of mouse embryo sections (Carpenter et al., 2004). BASP1 expression was also found in the nerve terminals of the peripheral nervous system (PNS) (Iino et al., 2004) and may therefore have functional roles within areas of the nervous system other than the brain. Evidence to support a role of BASP1 within the PNS is the involvement of the Wilms' tumour 1 (WT1) protein in optical and olfactory development; WT1 being an interacting protein of BASP1 (see section 1.5.7). Mouse embryos lacking both copies of *Wt1* (*Wt1*$^{-/-}$) have demonstrated a requirement for Wt1 in neuronal development of the retina and normal formation of the olfactory epithelium. These homozygous knockout embryos present with thinner than normal retinas, apoptosis of retinal ganglion cells, disruption of optic nerve fibre growth (Wagner et al., 2002a), thinner than normal sensory olfactory epithlium and fewer neuronal progenitor cells (Wagner et al., 2005).

Specifically in the mouse, polyclonal anti-BASP1 antibodies detected high expression of BASP1 within the CNS, lungs, heart, thymus, liver and tongue at developmental stage E15.5, the only developmental stage examined (Carpenter et al., 2004). Basp1 has also been implicated in the development of the mammary gland when it was shown to be upregulated in a microarray screen carried out to identify genes involved in migration and differentiation of terminal end buds (the specialised structures involved in mammary gland development) (Morris et al.,

2006). The expression of this gene was verified by RT-PCR, Western blotting and immunohistochemistry of mammary gland tissue sections.

Finally, BASP1 has been shown to be expressed within man using a human adult multitissue blot probed with anti-BASP1 antibodies (Carpenter et al., 2004). High expression was detected in the brain, lungs and testis, low expression in the kidney, spleen, ovary and heart, and no detectable expression in the liver or pancreas.

1.5.6 Roles of BASP1

The role of BASP1 within the nervous system was investigated by the use of mouse models. Following localised muscle paralysis (caused by either a subcutaneous injection of Botulinum toxin-A or a nerve crush), nerve sprouting and growth was observed, this regeneration being accompanied by induced re-expression of Basp1 in the intramuscular nerves (Caroni et al., 1997). To further elucidate the role of Basp1 in this regeneration process, transgenic mice were made using regulatory elements of the Thy-1.2 expression cassette (Caroni, 1997) to drive constitutive transgene expression of Basp1 in postnatal and adult neurons, the onset of expression being at a stage after early nervous system development, between approximately PD5-10 (Caroni et al., 1997). Four independent transgenic lines of the outbred Balb/C x C57BL/6 genetic background with high expression levels within the spinal motorneurons were analysed. A chicken BASP1 transgene was used so that both the endogenous and transgenic expression could be detected, due to the observation that antibody detection of BASP1 is species specific. Over-expression of Basp1 in the spinal motorneurons of these adult transgenic mice was found to induce spontaneous nerve sprouting at the neuromuscular junctions of all muscle types analyzed and enhanced nerve sprouting in response to localised muscle paralysis. This resulted in significant growth of synaptic structures as visualized on cryostatic tissue sections, with the nerves being visible by a combined silver esterase reaction. During this nerve regeneration, Basp1 expression was shown to be restricted to the neurons and to be caused by transgene expression with no reactivation of endogenous Basp1 expression. This inductive response was similar to that which

occurs in Gap43 over-expressing mice (Aigner et al., 1995), except that Basp1 over-expression induces fewer but longer nerve sprouts. Interestingly, double-transgenic mice over-expressing both proteins in the motoneurons caused enhanced nerve sprouting (particularly following localised muscle paralysis with Botulinum toxin-A), indicative of the proteins acting synergistically (Caroni et al., 1997). Indeed, during development, Basp1 and Gap43 are highly coexpressed within the nervous system in most neuron types (Frey et al., 2000). Within the adult however, expression of these two proteins is only partially overlapping in certain neuron types of the brain (Frey et al., 2000).

The role of Basp1 was further revealed by the study of a knockout mouse model of this gene. A targeting construct containing an NLS-*lacZ*-neo^r cassette (nuclear localization signal, β-galactosidase gene, neomycin resistance gene) was transferred into embryonic stem (ES) cells by homologous recombination, thereby replacing and hence inactivating one allele of *Basp1*. These ES cells were used to generate chimeric animals that were then bred to generate *Basp1*$^{+/-}$ and *Basp1*$^{-/-}$ null mice (Frey et al., 2000; Sadhu, 2006). Both *Basp1*$^{+/-}$ and *Basp1*$^{-/-}$ mice were born with the expected Mendelian frequencies, however *Basp1*$^{-/-}$ null mice displayed high postnatal mortality, with 66.2% dying between birth and PD120. Once reaching adulthood however, *Basp1*$^{-/-}$ null mice showed no elevated mortality compared to wildtypes. The reasons for the spontaneous postnatal lethality were not fully established, but severe abnormalities were found at the neuromuscular junctions as well as in the brain, specifically of the axonal and synaptic structures of the hippocampus and neocortex, and an enlargement of the ventricles. Of the surviving *Basp1*$^{-/-}$ mice, all showed signs of hyperactivity and adults were 50-70% the weight of their wildtype litter-mates although reasons for this difference in weight were not addressed in the study. *Basp1*$^{+/-}$ mice express approximately one third of wildtype *Basp1* levels as shown by northern analysis, have a less severe phenotype than full nulls, with some postnatal mortality and abnormalities at the neuromuscular junctions and average weights being ~80% of wildtype litter-mates. Basp1 has therefore been shown to be haplo-insufficient. The age at which the weights were measured was not mentioned, nor was an explanation as to the possible reasons for these differences in weight.

Studies carried out to determine the effect on nerve regeneration following localised muscle paralysis of *Basp1*$^{-/-}$ mice showed almost no stimulus-induced nerve sprouting (Frey et al., 2000). Interestingly, a complete rescue of this inability to form nerve sprouts was achieved by breeding onto a transgenic background with mice that overexpress either Basp1 (Caroni et al., 1997) or Gap43 (Aigner et al., 1995) in the adult motorneurons, a region with an overlapping expression pattern for these two proteins both during development and regeneration within the adult (Caroni et al., 1997). Cap23 and Gap43 may therefore act through the same regulatory pathway (Frey et al., 2000).

In comparison to mice devoid of Basp1 expression, *Gap43*$^{-/-}$ knockout mice displayed a different, more severe phenotype. These homozygous knockout mice were born with the expected Mendelian frequencies, however, around 45% died between PD0-P2 and another 45% died at weaning age between the second and third week after birth (Strittmatter et al., 1995). Although size and appearance at birth was normal, weight gain was poor over the first month, with average body weights being 55% that of wildtype and heterozygous littermates, all of which appeared phenotypically normal. The 10% of *Gap43*$^{-/-}$ mice that did survive past the third week continued to put on weight, albeit at a slower rate than the other littermates, and eventually caught up to 90% of normal weight at three months of age. Following the death of homozygous null mice, autopsy examination of the stomach and gut showed no presence of food, suggesting that mice were dying from starvation. *Gap43*$^{-/-}$ mice culled between E14-PD21 showed that body weights and organ weights were comparably diminished in size with no obvious abnormalities; organs examined included the liver, heart, lung, kidney, stomach and brain (Strittmatter et al., 1995). Developmental abnormalities found in *Gap43*$^{-/-}$ mice were disrupted axonal pathfinding of retinal ganglion cells at the optic chiasm of the retina, but by adulthood both the visual system and the nervous system as a whole was essentially normal, demonstrating that Gap43 does not appear to be a critical requirement for axonal outgrowth.

To further study the rescuing effects of Gap43 on mice lacking Basp1, *Basp1*$^{gap43/gap43}$ knockin mice were generated that expressed Gap43 in place of

Basp1 (Frey et al., 2000). These mice, apart from adult sterility (details not provided in the study) were in fact phenotypically almost entirely normal with no elevated mortality, differences in size or abnormalities at the neuromuscular junctions in the neocortex and hippocampus. Additionally, following localised muscle paralysis with Botulinum toxin-A these gene replacement mice showed nerve sprouting at the neuromuscular junctions, unlike Basp$^{-/-}$ null mice. These results therefore demonstrated that Gap43, although not sharing a high sequence homology with Basp1 (Maekawa et al., 1993; Wiederkehr et al., 1997), can functionally substitute for Basp1 *in vivo*, highlighting their similar functional properties.

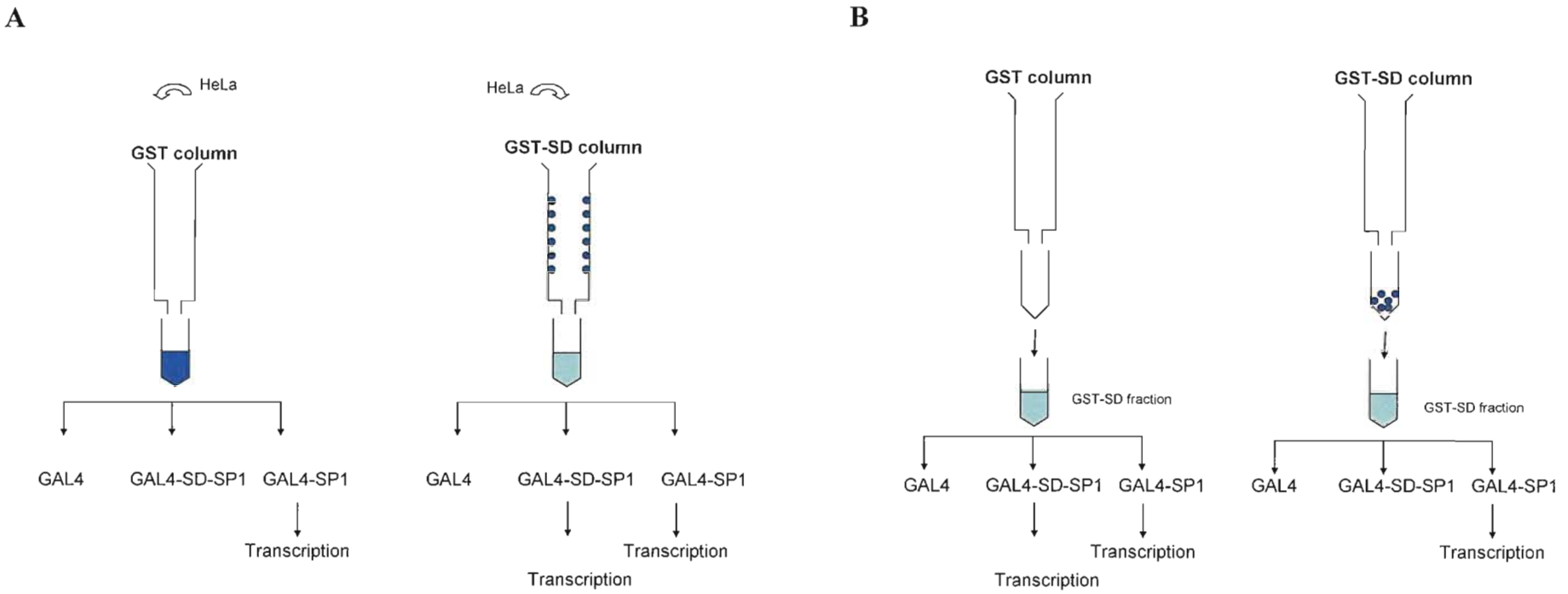
To further ascertain the synergistic functions of Basp1 and Gap43, different types of sensory neurons from the dorsal root ganglia of neonatal *Basp1*$^{-/-}$ mice and *Basp1*$^{gap43/gap43}$ knockin mice were cultured and patterns of neurite outgrowth were observed (Frey et al., 2000). It was shown that sensory neurons that usually express both Basp1 and Gap43 in wildtypes, grew abnormally in the absence of Basp1 (cultured from *Basp1*$^{-/-}$ mice) with extended axons that appeared thin and twisted, and growth cones frequently changing their direction of growth, as visualized by time-lapse video microscopy. The expression of Gap43 in place of Basp1 (cultured from *Basp1*$^{gap43/gap43}$ knockin mice) did not compensate for this abnormal growth in these neurons. Cultured sensory neurons that usually express only Basp1 in wildtypes also grew abnormally in the absence of Basp1, but in this example Gap43 could compensate for lack of Basp1 expression. The two proteins therefore have shared and unique roles in neurite outgrowth (Frey et al., 2000; Laux et al., 2000). Since Marcks as well as Basp1 and Gap43 has been shown to be involved in neurite outgrowth and synaptic plasticity through mouse knockout experiments (Laux et al., 2000; Stumpo et al., 1995), one could imagine an interesting study would be to test if Marcks have any rescuing effects on mice lacking Basp1, by generating a *Basp1*$^{Marcks/Marcks}$ knockin mice.

1.5.7 Identification of BASP1 as a WT1 interacting protein

Previous studies have shown that WT1 contains a suppression domain that can interact with a cofactor to mediate inhibition of a transcriptional activation

domain of WT1 (McKay et al., 1999; Wang et al., 1995). Carpenter et al. (2004) set out to determine the proteins involved in this functional inhibition by *in vitro* transcription assays and identified BASP1 as a component of the WT1 cosuppressor. These transcription assays were performed by passing HeLa nuclear extract through columns assembled with glutathione-s-transferase (GST)-linked agarose beads (GST only or GST fused with the WT1 suppression domain (SD)). The flow-through fractions were collected and then assayed with GAL4-fusion proteins (GAL4 only, GAL4 fused with the SP1 activation domain (GAL4-SP1), or GAL4 fused with SP1 and the WT1 suppression domain (GAL4-SD-SP1)) to test for their ability to drive expression of the synthetic G5E4T reporter construct (expression detected by primer extension) (Figure 1.5A). When the fractions from the GST only column were tested for transcriptional activation, like with untreated HeLa nuclear extract, expression of the reporter protein only occurred in the presence of the GAL4-SP1 fusion protein, showing that the suppression domain of WT1 can inhibit the activity of SP1. The fractions from the GST-SD column however showed expression of reporter protein in the presence of both the GAL4-SP1 and GAL4-SD-SP1 fusion proteins (Carpenter et al., 2004). The conclusion of these expression studies was that the free WT1 SD within the column must have depleted the nuclear extract of a cosuppressor thereby inactivating the suppression domain and preventing its inhibitory effect on SP1 during GAL4 expression studies. To test this hypothesis, the columns were washed extensively to remove any bound proteins and the eluates tested in transcription assay studies with the same GAL4 fusion proteins, by adding them to HeLa cell nuclear extract fractionated through the GST-SD column (Figure 1.5B). One of the eluates titrated from the GST-SD fraction showed an absence of reporter expression in the presence of GAL4-SD-SP1, demonstrating that the inhibitory function of the SD had therefore been restored. The proteins in these eluates were then separated by 2-dimensional gel electrophoresis with silver staining followed by identification using matrix-assisted laser desorption ionization-time-of-flight (MALDI-TOF) mass spectrometry. By comparison of the proteins present in the GST-SD eluates but not the GST eluates, at least one potential component of the cosuppressor was identified as BASP1.

Figure 1.5. The assays used to demonstrate and confirm that the WT1 suppression domain acts by recruiting a transcriptional cosuppressor (Carpenter et al., 2004). (A) HeLa nuclear extract was passed through columns assembled with glutathione-s-Transferase (GST)-linked agarose beads (GST only or GST fused with the WT1 suppression domain (SD)). The flow-through fractions were collected and then assayed with GAL4-fusion proteins (GAL4 only, GAL4 fused with the SP1 activation domain (GAL4-SP1), or GAL4 fused with SP1 and the WT1 suppression domain (GAL4-SD-SP1)) to test for their ability to drive expression of the synthetic G5E4T reporter construct. Reporter expression in the presence of GAL4-SD-SP1 with HeLa nuclear extract which has been passed through the GST-SD column shows that a cosuppressor that usually prevents WT1 transcriptional activity has been depleted from the extract and remains in the column. (B) After the initial assay (see Figure) the GST column and the GST-SD column were washed extensively to remove any bound proteins. The eluates were then tested in transcription assays with the GAL4-fusion proteins (GAL4 only, GAL4 fused with the SP1 activation domain (GAL4-SP1), or GAL4 fused with SP1 and the WT1 suppression domain (GAL4-SD-SP1)) to test for their ability to drive expression of the synthetic G5E4T reporter construct. These GAL4 fusion proteins were added to HeLa nuclear extract which had been fractioned through the GST-SD column. An eluate titrated from the GST-SD column when added to the GST-SD fraction showed an absence of reporter expression in the presence of GAL4-SD-SP1. This absence of activity demonstrates that the eluate contains a cosuppressor titrated from the column that is required for inhibition of transcriptional activity by the WT1 SD.


A
HeLa
GST column
GAL4
GAL4-SD-SP1
GAL4-SP1
Transcription
HeLa
GST-SD column
GAL4
GAL4-SD-SP1
GAL4-SP1
Transcription
Transcription
B
GST column
GST-SD fraction
GAL4
GAL4-SD-SP1
GAL4-SP1
Transcription
Transcription
GST-SD column
GST-SD fraction
GAL4
GAL4-SD-SP1
GAL4-SP1
Transcription

To verify that BASP1 interacts directly with the WT1 SD, co-immunoprecipitation assays were carried out (Carpenter et al., 2004). GST and GST-BASP1 were linked to glutathione agarose beads and placed in the presence of the three GAL4 fusion proteins. Using anti-GAL4 antibodies, GAL4-SD-SP1 but not GAL4-SP1 was shown to interact with GST-BASP1, suggesting that binding to BASP1 is mediated via the WT1 suppression domain.

Another functional assay to test the transcriptional activation function of BASP1 involved the use of RNAi (Carpenter et al., 2004). A luciferase reporter under the control of the amphiregulin gene promoter (a WT1 target) was first transfected into HeLa cells in the presence or absence of a construct in which CMV regulatory elements were used to drive WT1 expression. WT1 does evoke some activation of the amphiregulin promoter in this system as observed by low levels of luciferase activity. However, when the HeLa cells were cotransfected with a construct designed to produce siRNAs to specifically target and degrade *BASP1* RNA, a high level of luciferase activity in the presence of WT1 was observed, indicating potent induction of the amphiregulin promoter. These studies thereby show conclusively that endogenous BASP1 can silence WT1 transcriptional activation function.

As BASP1 is a cosuppressor of WT1, a tumour suppressor gene for Wilms' tumour of the kidney, expression studies were carried out to analyze the expression and localization of BASP1 within the kidney (Carpenter et al., 2004). Immunofluorescence of mouse kidney M15 cells and human paediatric kidney tumour G-401 cell detected endogenous BASP1 entirely within the nucleus, whereas endogenous BASP1 was detected in the nucleus and in the cytoplasm of monkey kidney Cos-1 cells. M15 cell nuclear extracts were prepared and immunoprecipitated with anti-WT1 or anti-BASP1 antibodies. It was shown that anti-WT1 coprecipitates both WT1 and a 40kDa form of BASP1 ($BASP1_{40}$) (as detected by SDS-PAGE on 8-18% gradient acrylamide) but not a larger myristoylated form of BASP1. Anti-BASP1 coprecipitates BASP1 (but only if BASP1 is not myristoylated) and WT1, demonstrating that the two endogenous proteins can associate within M15 cells and suggesting unique and distinct cellular roles for BASP1 dependent upon its myristoylation status. Recently, it

has been shown that both WT1 and BASP1 are expressed primarily in the nucleus and partially colocalize within mouse MPC1 cells (Mundel et al., 1997), a podocyte precursor cell line (Stefan Roberts, personal communication).

In the developing mouse kidney Basp1 protein detected by immunohistochemistry has been shown to localise to the forming nephron structures at E15.5. In the adult mouse and human kidney, Basp1 was restricted to the specialised podocyte cells and has been shown to localize within the nucleus and the cytoplasm of these cells. This pattern of Basp1 protein expression within the kidney recapitulates WT1 protein expression in this organ, which led Carpenter et al. (2004) to postulate that these two proteins could act through the same pathway involved in the control of kidney development and the maintenance of adult kidney function.

With this possible kidney development role in mind, the WT1/BASP1$_{40}$ complex was analyzed in the podocyte precursor cell line (MPC1) before and after differentiation (approximately four days following induction of differentiation). Both proteins were analyzed for DNA binding partners using ChIP to ascertain which gene promoters of the previously proposed WT1 target genes, if any, they bound to. RT-PCR was then used to quantify the relative levels of RNA from these genes. Among others it was found that the protein complex colocalizes with and regulates the genes podocalyxin, nephrin, bak and c-myc2. Interestingly, the WT1/BASP1$_{40}$ complex was not seen to colocalize at the amphiregulin promoter, another previously proposed WT1 target gene. Although this does not preclude the presence of these proteins at the amphiregulin promoter, as other factors present at the promoter may be able to block the antibody binding sites leading to a failure of the ChIP assay, the fact that neither protein was shown to colocalize with the amphiregulin promoter makes this less likely (Stefan Roberts, personal communication). This result appears to conflict with an RNAi functional assay (Carpenter et al., 2004) which shows that when BASP1 is downregulated in the presence of a WT1 expressing construct driven by CMV, a higher level of luciferase activity under the control of the amphiregulin promoter was observed (for further details see above). Although this is an artificial system and therefore not as physiologically relevant as *in vivo*

studies, it demonstrates that at least under some circumstances, BASP1 is able to exert an effect on the ability of WT1 to control amphiregulin. If this control of amphiregulin by WT1 and BASP1 also occurs in podocyte cells, it is possible that these two proteins may act indirectly to control amphiregulin, due to their absence at the promoter as shown by ChIP. Points to bear in mind are that the podocyte cells used for the ChIP assay were grown *in vitro*, so could differ from endogenous podocyte cells; also, it is possible that the exact region of the amphiregulin promoter used for the ChIP assay and RNAi assay may have been different; finally, it is possible that WT1 and BASP1 bind to another regulatory region of amphiregulin in order to exert control over this gene.

Looking in more detail at the binding of the WT1/BASP1$_{40}$ complex to the promoter of podocalyxin, WT1 was shown to be robustly bound to the promoter both before and after differentiation, with an increased efficiency of promoter occupancy occurring at three days following differentiation. BASP1 however was only robustly bound before and shortly after the onset of differentiation, with an increased efficiency of promoter occupancy at three days following differentiation as with WT1; by five days following the induction of differentiation binding has significantly reduced. Looking at the RNA expression levels of these genes by RT-PCR, podocalyxin levels dramatically decreased initially, then levels recovered over several days of differentiation; those of WT1 decreased slightly but then returned to the same levels as before differentiation; and those of BASP1 increased slightly initially but were then reduced. The recovery of podocalyxin expression therefore coincides with the reduction of bound BASP1 at its promoter, consistent with the cosuppressor function proposed for BASP1.

Interestingly, the reduction of BASP1$_{40}$ coincided with an increase in BASP1$_{60}$ isoforms within the differentiated cells as shown by immunoblotting of cell fractions. BASP1 contains two SUMO-modification sites and *ex vivo* studies showing that BASP1 was able to bind to SUMO-3 (Stefan Roberts, article in preparation) lead to the prediction that this increase from a BASP1$_{40}$ species to a BASP1$_{60}$ species could be due to sumoylation. Documented observations that sumoylation can directly effect transcriptional regulation and can cause

relocation within a cell (Gill, 2004) lead to this possibility being addressed as a mechanism by which $BASP1_{40}$ is redirected away from the chromatin. MPC1 fractions were taken before and after differentiation, to identify the cellular location of $BASP_{40}$. It was found that $BASP1_{40}$ was present in the chromatin fraction before differentiation, indicative of a role in transcription, but not found in any fraction after differentiation. $BASP1_{60}$ however was found in the matrix fraction after differentiation, demonstrating that sumoylation could be the mechanism allowing relocalization of this protein in order to direct BASP1 to the nuclear matrix, away from the chromatin, and therefore away from its role in transcription. Together with the elevated levels of podocalyxin, this fits with previous observations demonstrating the BASP1 can act as a cosuppressor of WT1 by binding to its suppression domain and blocking the transcriptional activation domain of WT1; a lifting of this suppression therefore allowing transcriptional activation to occur (Carpenter et al., 2004).

Aims of Thesis

The transcriptional activation domain of WT1 is subject to regulation by an N-terminal suppression domain that has been proposed to function by recruiting WT1 cosuppressor proteins that can block the function of the transcriptional activation domain. Transfection assays have identified Brain acidic soluble protein 1, BASP1, as a component of the WT1 cosuppressor that can regulate WT1 transcriptional activity. BASP1 is a nuclear factor that associates with WT1 in cells that naturally express both proteins. Analysis of a limited number of embryonic and adult kidney sections has revealed that BASP1, coincident with WT1, is present in the developing nephron structures of the embryonic kidney and its expression is restricted to the highly specialized podocyte cells of the adult kidney. BASP1 however is also more widely expressed than WT1.

Another factor thought to be important in the regulation of WT1 is the WT1 antisense transcript, *WT1-AS*. *WT1* and *WT1-AS* RNA transcripts, as well as WT1 protein are coordinately expressed in the developing nephron structures of the embryonic kidney and the podocyte cells of the adult kidney. The promoter region of the *WT1-AS* transcript has been shown to be transcriptionally activated by WT1 and it also appears that *WT1* and *WT1-AS* may reciprocally regulate each others expression. A region of this antisense transcript termed Mouse Conserved Region II, MCRII, has a shared 94% sequence identity between mouse and man, strongly suggesting a regulatory role for this element.

The first aim of this thesis (chapter 3) was to ascertain within the mouse, molecular and genetic information on BASP1 and to further investigate the expression pattern of this gene. The second aim (chapter 4) was to generate a knockout mouse model of BASP1 in order to investigate the effect on kidney development and development as a whole. The third aim (chapter 5) was to ascertain within the mouse, molecular and genetic information on *WT1-AS* and to generate a mouse model lacking the Wt1 antisense regulatory region, MCRII, in order to investigate the effect on kidney development.

Materials and Methods

2.1 Molecular Protocols

(Sambrook et al., 1989)

2.1.1 Restriction digests

Digestion of DNA was carried out with restriction enzymes (Promega or NEB) following the manufacturer's instructions. The reactions were set up in 1.5ml microcentrifuge tubes containing DNA, $^{1}/_{10}$ volume of restriction enzyme buffer, 100µg/ml BSA, 5U of enzyme per µg of DNA and water to the required volume (ensuring the volume of enzyme never exceeded 10% of the total reaction volume). Most enzymes were incubated at 37°C for 1-2 hours (some enzymes required different incubation temperatures). When required, restriction enzymes were heat inactivated at 65°C for 15mins.

2.1.2 Agarose gel electrophoresis

Agarose gels were made using powdered agarose (Invitrogen). A 1% weight per volume gel was made by adding agarose to 1x TAE (0.04M Tris-acetate, 0.001M EDTA). Before pouring into a gel casting tray, 500ng of ethidium bromide (10mg/ml) per ml of agarose was added to aid visualization of nucleic acids. Loading dye was added (Promega 6x, 1.7µl per 10µl of solution) to each DNA/RNA sample before loading into the wells of the gel. Samples were electrophoresed at 90-100V for 45mins or until the marker ladder was nicely separated. The DNA/RNA samples were visualized under ultra-violet light. When RNA samples were to be run, the gel tank, gel tray and combs were cleaned first with 3% hydrogen peroxide for 30mins, then rinse with sterile diethyl pyrocarbonate (DEPC)-treated water (1ml DEPC (Sigma) per 1 litre water). RNase-free agarose gels were made using 1x TAE made with DEPC-treated water.

2.1.3 Preparation of plasmid vector DNA and insert DNA for ligation

DNA to be ligated together was cut with appropriate restriction enzymes and then heat inactivated. Insert DNA was purified from a 0.8-1% agarose gel using 'Wizard® SV gel clean-up system' (Promega), following the manufacturer's instructions. When vector DNA was cut with a single restriction enzyme, 2 units of Shrimp Alkaline Phosphate (USB) was used to dephosphorylate the free DNA ends after completion of digestion. When necessary, DNA generated by restriction enzymes that give 5' or 3' overhangs were rendered blunt either with Klenow (Promega), used to fill-in 5' protruding ends, or T4 DNA Polymerase (Promega), used to remove 3' single-strand overhangs by its 3'-5' exonuclease activity. In each case, reaction conditions were as per the manufacturer's instructions.

2.1.4 Annealing Oligonucleotides

Two oligonucleotides (100pmole/µl), 1x annealing buffer (10mM Tris-HCl pH7.5, 100mM NaCl, 1mM EDTA) and DEPC-treated water were incubated at 65°C for 10 mins then allowed to cool slowly at room temperature for 1-2 hours. These double stranded linkers were then used in ligation reactions. Linkers were made with overhang or blunt ends to allow ligation to compatible vector DNA ends. Vector DNA was never dephosphorylated as oligonucleotides lack 5' phosphates.

2.1.5 Ligation of DNA

50-200ng of digested vector DNA was added together with a 3-5x molar excess of insert DNA, ½ volume of ligation buffer (Promega) and 10U T4 DNA ligase (Promega), then incubated overnight at 4°C. Ligations were stored at –20°C until needed.

2.1.6 Bacterial transformations

Ligation reactions were microdialysed on nitrocellulose filters (Millipore) floating in water, for half an hour to lower the salt concentration. The entire ligation reaction was then added to 40µl of electrocompetent cells (section 2.1.7)

in a 1mm-gap electroporation cuvette (BTX) placed on ice. A single electrical pulse (1.25kV, 25μF, 200Ω) was then applied to the cells and DNA in the cuvette. 1ml of cold LB medium was then added to the cuvette before transferring the cells and broth to a 15ml snap-cap tube (Falcon) and incubating at 37°C for 1hour. After the 1hour recovery period, 50μl of culture was spread onto several LB-agar plates containing the appropriate antibiotic (50μg/ml carbenicillin or 25μg/ml kanamycin; Sigma), and grown overnight at 37°C. Occasionally XL10-Gold® ultracompetent cells (Stratagene) or XL1-Blue supercompetent cells (Stratagene) were used, following manufacturers instructions.

2.1.7 Making electrocompetent *E.coli* cells

A glycerol stock of DH5α *E.coli* cells was streaked onto a freshly prepared LB agar (Sigma) plate. One colony was picked and grown overnight in 10ml LB medium (Sigma). The following day, 1 litre of LB medium was inoculated with 10ml of the overnight culture. Cells were then grown at 37°C in a shaking incubator for 2-4hours until the OD_{600} was between 0.5-0.8. The cells were then chilled on ice for 15mins (the cells were kept cold from this point on by keeping them on ice in the cold room and pre-chilling the centrifuge and rotors to 4°C). The chilled cells were harvested by centrifuging for 10mins at 1300rcf. The supernatant was discarded and the pelleted cells resuspended by pipetting and vortexing in 500ml ice cold sterile milliQ water (25ml was added initially to aid resuspension of cells, before the final volume of water was added). Cells were then centrifuge again for 10mins at 1300rcf. The supernatant was discarded and the pelleted cells resuspended in 200ml ice cold sterile milliQ water, then centrifuged for 10mins at 1800rcf. The supernatant was discarded again and the pelleted cells resuspended in 100ml ice cold sterile 10% glycerol, then centrifuged for 10mins at 3000rcf. The supernatant was poured off carefully and the pelleted cells resuspended in 2ml of ice cold sterile 10% glycerol. Cells were centrifuge for a final time for 4mins at 10,000rcf. The supernatant was pipetted off and 10% glycerol was added to a total volume of 3ml. This suspension was then aliquoted into 0.5ml microcentrifuge tubes, 40μl per tube, snap frozen in liquid nitrogen, then store at –80°C.

2.1.8 E.coli plasmid small scale preparation

Miller and Miller method, University of Oregon

Individual colonies were picked with a sterile toothpick from LB-agar transfection plates and grown in ~4ml LB broth containing the appropriate selective antibiotic, overnight at 37°C in a shaking incubator. 1.5ml of the overnight culture was transferred to a microcentrifuge tube and centrifuged for 30sec at 10,000rcf. The supernatant was discarded and the pellet resuspended in 100µl of cold Solution 1 (50mM Glucose, 25mM Tris HCl pH8, 10mM EDTA pH8, 10g/µml RNase; filter sterilised). The samples were vortexed until no cell clumps remained. 200µl of Solution 2 (0.2N NaOH, 1% SDS, water) was added and the tubes inverted 4-6 times to mix. 150µl of Solution 3 (5M Potassium acetate, 11.5% Glacial acetic acid; autoclaved) was added and the tubes inverted 4-6 times to mix. The samples were then centrifuged for 2mins at 10,000rcf. 1µl of DEPC was added and vortexed before incubating at 70°C for 10mins. The samples were cooled on ice for 10mins before centrifuging for 2mins at 10,000rcf. The supernatants were then transferred to fresh 1.5ml microcentrifuge tubes. 900µl of 100% ethanol was added to the supernatants, mixed well, then centrifuged for 15mins at 10,000rcf. The ethanol was poured off, and the pellets washed in a small amount of 100% ethanol. Pellets were left to air-dry before finally being resuspended in 55µl TE buffer.

2.1.9 *E.coli* plasmid small scale preparation

Wizard® Plus SV Miniprep DNA Purification System (Promega)

Individual colonies were picked with a sterile toothpick from LB agar transfection plates and grown in ~4ml LB broth containing the appropriate selective antibiotic, overnight at 37°C in a shaking incubator. 1.5ml of the overnight culture was transferred to a microcentrifuge tube and centrifuged for 30sec at 10,000rcf. The supernatants were discarded and the pellets resuspended in 250 µl Wizard Plus Cell Resuspension Solution. 250 µl of Wizard Plus Lysis Solution was added and mixed by inverting the tubes. 10 µl of Alkaline Protease Solution was then added and the tubes inverted again. After a 5 minute

incubation at room temperature, 350 µl of Wizard Plus Neutralization Solution was added and after inverting to mix, the samples were centrifuged for 10mins at 10,000rcf. The cleared lysates were transferred into Wizard Spin Columns and centrifuged for 1min at 10,000rcf. The flow-through was discarded. After addition of 750 µl of Wizard Plus Wash Solution, samples were centrifuged again for 1min at 10,000rcf. The wash step was repeated with 250 µl Wash Solution. The columns were then transferred to fresh 1.5 ml microcentrifuge tubes and 100µl of Nuclease-Free Water was added to each sample. After a final spin for 1min at 10,000rcf the columns were discarded and the DNA stored at -20°C.

2.1.10 DNA sequencing

DNA sequencing was performed either in the 'DNA Sequencing Facility, University of Bath' or by 'MWG' using an ABI DNA sequencer module. 500-1000ng of DNA was used with 10µM sequencing primer. Sequences were analyzed using the 'BioEdit Software package'.

2.1.11 Glycerol stocks

Bacterial colonies containing a required plasmid were frozen for archival storage. 0.7ml of liquid bacterial culture and 0.3ml of 50% sterile glycerol were gently mixed in a 1.5ml microcentrifuge tube and frozen at –80°C.

2.1.12 Southern blotting

DNA samples were digested with the required restriction enzyme. DNA samples were then loaded onto an agarose gel and gel electrophoresis was carried out (sections 2.1.1 and 2.1.2). Ethidium bromide stained agarose gels were visualized under UV light and photographs were taken with a ruler alongside the DNA ladder. Gels were then placed in denaturation solution (0.5M NaOH, 1M NaCl) for 30 minutes with gentle shaking. The gel was transferred into neutralization solution (1.5M NaCl, 50mM Tris-HCl, 1mM EDTA pH8, pH 7.2) for 15 minutes with gentle shaking. The neutralization step was repeated. A capillary blot was

set up as described in Sambrook et al. (1989). A Nylon transfer membrane (Osmonics) was used for the capillary blot, and 20x SSC was used as the blotting buffer (3M NaCl, 0.3M sodium citrate, pH7.0). Capillary transfer was allowed to proceed overnight. The following day the membrane was rinsed in 2x SSC to remove any adhering agarose. The nylon membrane was allowed to air dry for 1 hour then UV crosslink in a CL-1000 Ultraviolet Crosslinker (UVP) at setting 24000μJ/cm^2.

2.1.13 Southern hybridization

UV crosslinked nylon membranes were rolled up (DNA on the inner surface) and placed in a hybridization bottle. 15ml of hybridization buffer (0.5M NaH_2PO_4, 1mM EDTA, 7% SDS, 1% BSA) and 150μl salmon sperm (Sigma type III sodium salt, 10mg/ml denatured and sheared, boiled for 5mins immediately before use) was added to the bottle and the membrane was allowed to pre-hybridize for 2 hours in a rotating 65°C incubator.

Radiolabelled probes were made by adding ~25ng linearised probe DNA (made up to 13μl with water), 4μl HighPrime (Roche; 1U/μl klenow polymerase, 0.125mM dATP, 0.125 mM dGTP, 0.125 mM dTTP in 50% (v/v) glycerol) and 30μCi radiolabel α-^{32}P dCTP (Amersham) to a 1.5ml microcentrifuge tube, then heating for 20 minutes at 37°C. TE was added to the hybridized probe to a final volume of 100μl. Unincorporated nucleotides were separated from the radiolabelled probe by the spin-column method, as follows: A 1ml syringe was plugged with polyallomer wool then filled with reconstituted Sephadex G50 (Sigma). The syringe was placed in a 15ml tube, centrifuged for 2mins at 1000rcf to equilibrate, and the flow-through was discarded. The 100μl of labelling reaction was added to the spin-column then centrifuged for 2mins at 1000rcf, with the flow-through collected in a 1.5ml microcentrifuge tube. A further 100μl of TE was then added to the column and a final spin was carried out, again collecting the flow-through. The two flow-through fractions containing the hybridized probe were transferred to a fresh 1.5ml microcentrifuge tube, then boiled for 4-5 minutes immediately prior to addition to the hybridization bottle.

After the 2 hour pre-hybridization period, the hybridization buffer was poured away and fresh buffer, boiled salmon sperm and the hybridized probe was added to the hybridization bottle. The blot was left to hybridize overnight. The next day, the hybridization buffer was poured off. The bottle was swilled with ~50ml wash solution 1 (0.5% BSA, 5% SDS, 40mM NaH_2PO_4, 1mM EDTA) heated to 65°C, then poured off. The bottle was then filled $^1/_4$ full with heated wash solution 1 and incubated in a rotating 65°C incubator for 15 minutes. This wash step was then repeated. The wash solution 1 was poured off and the bottle filled ¾ full with heated wash solution 2 (1% SDS, 40mM NaH_2PO_4, 1mM EDTA) and incubated at 65°C for 1 hour. This wash step was then repeated. After these wash steps, carried out to remove any non-specific binding of the probe, the hybridized membrane was wrapped in ™SaranWrap and placed DNA side up in a cassette with GRI AX film (GRI) placed on top. The cassette was stored at –80°C and processed in a dark room, through an X-OGRAPH X-ray processor several days later.

2.1.14 Grunstein colony hybridization

Four pieces of filter paper were laid on top of four separate lengths of cling film and saturated with 10% SDS, denaturation solution (0.5M NaOH, 1M NaCl), neutralization solution (1.5M NaCl, 50mM Tris-HCl, 1mM EDTA pH8, pH 7.2) and 2x SSC (3M NaCl, 0.3M sodium citrate, pH7.0) respectively. A circular Hybond nylon membrane was gently laid on top of colonies of an agar plate for 1 minute and using a hypodermic needle three holes were made in the membrane and agar plate to serve as an orientation cue. The membrane was then transferred to the filter paper saturated in 10% SDS, colony side up, for about 3 minutes in order to wet the membrane. It was then transferred to the filter paper saturated with denaturation solution followed by neutralization solution for 5 minutes each. Finally the membrane was transferred to the filter paper saturated with 2x SSC, rubbed gently by hand then placed in a container filled with 2x SSC on a rocker for several minutes. This procedure was repeated with a second membrane then the agar plate was placed at 37°C overnight to allow the colonies to re-grow. The membranes were allowed to air dry for 1 hour then UV crosslink in a CL-

1000 Ultraviolet Crosslinker (UVP) at setting 60,000μJ/cm^2. Radioactive labelling of probes with P^{32} was carried out as in the Southern hybridization protocol (section 2.1.13) but with a 1 hour 37°C labelling reaction time. After overnight hybridization at 65°C the membranes were washed twice for 20 minutes at 65°C with a wash solution containing 1x SSC and 1% SDS, followed by washing twice for 20 minutes at 65°C in a wash solution containing 0.1x SSC and 0.1% SDS if the membrane signal was still very strong after the first washes. The membranes were wrapped in ™SaranWrap and placed DNA side up in a cassette with autoradiographic, AX film (GRI) placed on top. After 15-60mins at room temperature the film was processed in a dark room through an X-OGRAPH X-ray processor. Colonies that showed hybridization with the probe on both duplicate membranes were then grown in order to extract plasmid DNA.

2.1.15 Northern blotting

Gel apparatus was treated with 3% hydrogen peroxide to get rid of contaminating RNases. A 1% (w/v) ethidium bromide-free gel was made by adding agarose to RNase-free 1x MOPS buffer (0.02M MOPS pH7.0, 2mM sodium acetate, 1mM EDTA pH8.0), 17% formaldehyde and DEPC-treated water. 25μg of total RNA, 2% formamide, 1x MOPS buffer and 6.25% formaldehyde was heated to 65°C for 5 mins, cooled on ice and 1x loading buffer was added before loading into a well of the gel. Gel electrophoresis was carried out in running buffer (1x MOPS buffer, DEPC-treated water) for ~4 hours at 5V/cm length of gel. Fermentas High Range RNA ladder was used (following manufacturers instructions) with its own loading buffer. Following electrophoresis the RNA ladder was stained separately for 1 hour in 1x TAE buffer containing 500ng of ethidium bromide (10mg/ml) per ml of buffer in order to visualize the separate RNA bands under UV light and for photographs to be taken alongside a ruler. The gel was rinsed briefly in DEPC-treated water before blotting. A capillary blot was set up as described in Sambrook et al. (1989). A Nylon transfer membrane (Osmonics) was used for the capillary blot, and 20x SSC was used as the blotting buffer (3M NaCl, 0.3M sodium citrate, pH7.0). Capillary transfer was allowed to proceed overnight. The following day the membrane was allowed to air dry then UV crosslink in a CL-1000 Ultraviolet Crosslinker (UVP) at setting 24000μJ/cm^2.

2.1.16 Northern hybridization with DNA probes

UV crosslinked nylon membranes were rolled up (RNA on the inner surface) and placed in a hybridization bottle. Follow the Southern hybridization protocol was followed.

2.1.17 Northern hybridization with RNA probes

UV crosslinked nylon membranes were rolled up (RNA on the inner surface) and placed in a hybridization bottle. 50ml of hybridization buffer (6x SSC, 5x Denhardt's (0.5g ficoll, 0.5g polyvinylpyrrolidone, 0.5g BSA), 0.5% SDS, 50% formamide), 15µg yeast tRNA (preheated to 95°C for 5 mins) and 15µg herring sperm (preheated to 95°C for 5 mins) was added to the bottle and the membrane was allowed to pre-hybridize for 2 hours in a rotating 42°C incubator.

Radiolabelled probes were made by adding together, to a total of 10µl with water, linearized probe DNA (0.2pmoles), DTT (10mM), rNTPs (rATP, rCTP, rGTP 0.5mM each), 1x buffer (40mM Tris-Cl pH7.5, 6mM $MgCl_2$, 5mM NaCl), RNase inhibitor (10U), BSA (0.1µg), radiolabel α-^{32}P rUTP (50µCi, Amersham), T3 or T7 RNA polymerase (20U, Boehringer Mannheim), then heating for 1-2 hours at 37°C followed by the addition of DNase (1µl) for 15 mins at 37°C. The RNA probe was then purified away from the unincorporated nucleotides using an RNeasy MinElute cleanup kit (Qiagen) following manufacturer's instructions before its addition to the hybridization bottles. The blot was left to hybridize overnight. The next day, the hybridization buffer was poured off and the blot washed once with wash solution A (2x SSC, 0.5% SDS) for 5 mins at room temperature, once with wash solution B (2x SSC, 0.1% SDS) for 15 mins at 42°C, twice with wash solution C (0.1x SSC, 0.5% SDS) for 1 hour at 68°C and once in wash solution D (0.1x SSC, 0.1% SDS) for 30 mins at 68°C. After these wash steps carried out to remove any non-specific binding of the probe, the hybridized membrane was wrapped in ™SaranWrap and placed RNA side up in a cassette with GRI AX film (GRI) placed on top. The cassette was stored at –80°C

and processed in a dark room, through an X-OGRAPH X-ray processor several days later.

2.1.18 Coomassie stained SDS-PAGE urine gels

Urine samples were allowed to completely thaw on ice before transferring 5µl into a fresh 1.5ml microcentrifuge tube containing 5µl of 2x reducing buffer (250mM Tris-HCl pH6.8, 2% SDS, 10% glycerol, 10% β-mercaptoethanol, 0.02% bromophenol blue). The urine samples in reducing buffer were then boiled for 10 mins followed by immediate cooling on ice. Boiled sample were then loaded into wells of a precast 10% polyacrylamide Tris-HCl resolving gel (Criterion, BIO-RAD), in a Criterion™ cell midi tank (BIO-RAD) with 1x running buffer (0.12M Tris, 0.96M glycine, 0.02M SDS) and run at 180V for 1 hour. 10µg of boiled BSA in reducing buffer was used as a positive control. 10µl of precision plus protein dual color standards ladder was used as a marker (BIO-RAD). Following electrophoresis the gel was removed from its plastic casing and placed into coomassie stain solution (40% methanol, 10% acetic acid, 3mM Brilliant Blue R (Sigma)) on a shaker for 1 hour. The gel was then washed in coomassie destain solution (40% methanol, 10% acetic acid) for 15 mins followed by two 1 hour washes. Photographic images were taken under transilluminator white light set at 20-80 milliseconds on an Alphaimager™ 3400 (GRI).

2.1.19 cDNA synthesis

A 'SuperScript II Reverse Transcriptase Kit' (Invitrogen) was used for cDNA synthesis. 1µg of RNA was added to 12µl sterile DEPC-treated milliQ water, 1µl dNTP mix (10mM each) and either oligo(dT) primers 500µg/ml / random primers 50-250ng / or gene specific primers 2pmole. RNA and primers were heated to 65°C for 5mins. 4µl (5x) first-strand buffer, 2µl 0.1M DTT, and 1µl water was added to the tube which was then incubated for 2mins at 42°C when either oligodT or gene-specific primers were used or at 25°C when random primers were used. 1µl (200units) of SuperScript II Reverse Transcriptase (RT) was added to RT+ tubes (1µl of water was added to RT- negative control tubes).

The RNA was then incubated at 42°C for 50mins. The reaction was inactivated by heating for 15mins at 70°C. 1µl of cDNA was used for each RT-PCR reaction.

2.1.20 PCR/RT-PCR

1µl of template (e.g. 0.1-1µg purified genomic DNA, 5-50pg plasmid DNA, cDNA, bacterial colony lysates), 1µl forward primer (6µM), 1µl reverse primer (6µM), 10µl (2x) master mix (Promega), and 7µl nuclease-free water were added together in a 0.2ml microcentrifuge tube. PCRs were carried out in a TouchGene cycler (Techne), with an initial denaturation at 94°C for 3mins, then 30-40 cycles of denaturation (94°C 30-60secs), annealing (50°C-68°C 30-60secs; 5°C below the average Tm of the two primers), and extension (72°C 30-60secs), followed by a final polymerase extension at 72°C 5mins. PCR primer pairs are listed within the methods section of each results chapter.

2.1.21 cDNA synthesis to reduce endogenous priming

An 'EndoFree™ kit' (Ambion) was used for cDNA synthesis. ~150ng of poly(A)$^+$ RNA and 10pmole of a gene specific primer was added to sterile DEPC-treated milliQ water up to a total of 7µl and incubated for 5 mins at 70°C followed by 5 mins at 49°C. 12µl of pre-mixed reagents, pre-warmed for 5 mins at 49°C was then added to each of the reaction tubes, taking care not to allow the reaction to drop much below 49°C in the process. Premixed reagents consisted of 2µl (10x) RT buffer, 8µl dNTPs (2.5mM each), 1µl RNase inhibitor (10U) and 1µl oligo(dA)$_{80}$ (1µg). 1µl of EndoFree RT was then added to RT+ tubes (1µl of water was added to RT- negative control tubes). The reaction was then incubated at 49°C for 2 hours followed by inactivation by heating for 45 mins at 80°C. 1µl of cDNA was used for each RT-PCR reaction.

2.1.22 3' rapid amplification of cDNA ends (3'-RACE)

The SMART™ RACE cDNA Amplification system (Clontech) was used for 3'-RACE reactions. ~1µg total RNA and 1µl 3'-CDS primer (10µM) in a total volume of 5µl deionized water was added to a 0.2ml microcentrifuge tube and heated for 2 mins at 70°C followed by cooling on ice for 2 mins. 2µl first-strand

synthesis buffer (5x), 1µl DTT (20mM), 1µl dNTP mix (10mM each) and 1µl Superscript II reverse transcriptase (200units/µl, Invitrogen) were then added to the tube and incubated for 1.5 hours at 42°C. The first-strand synthesis cDNA reaction product in the PCR tube was diluted by adding 100µl Tricine-EDTA buffer (10mM Tricine-KOH pH8.5, 1mM EDTA) and heated for 7 mins at 72°C. The Advantage® 2 PCR enzyme system (Clontech) was then used to carry out RT-PCR reactions on the synthesised cDNA. 2.5µl 10x Advantage® 2 PCR buffer (400mM Tricine-KOH pH8.7, 150mM KOAc, 35mM $Mg(OAc)_2$, 37.5µg/ml BSA, 0.05% Tween-20, 0.05% Nonidet-P40), 0.5µl dNTP mix (10mM each), 0.5µl 50x Advantage® 2 polymerase mix (50% glycerol, 15mM Tris-HCL, 75mM KCl, 0.05mM EDTA) and 17.25µl nuclease-free water was added to a PCR tube and vortexed. Next, 1.25µl 3-RACE cDNA from the first-strand synthesis reaction, 2.5µl 10x Universal Primer Mix and 0.5µl gene specific primer (10µM) was added to a PCR tube to carry out RT-PCR. RT-PCR reactions were carried out in a TouchGene cycler (Techne), with 35 cycles of denaturation (94°C 10secs), annealing (68°C 10secs) and extension (72°C 2 mins). The RT-PCR product was then used to seed a nested RT-PCR reaction. 2µl of the first RT-PCR product, 2.5µl 10x Advantage® 2 PCR buffer, 0.5µl dNTP mix (10mM each), 0.5µl 50x Advantage® 2 polymerase mix, 1µl Nested Universal Primer (10µM), 1µl nested gene specific primer and 17.5µl nuclease-free water was added to a 0.2ml microcentrifuge tube and a second RT-PCR reaction was carried out in a TouchGene cycler, with 35 cycles of denaturation (94°C 10secs), annealing (68°C 10secs) and extension (72°C 2 mins).

2.1.23 Long template PCR

'Advantage®2 PCR enzyme system' (Clontech) was used for long template PCR. Template DNA needed to be of high quality (high molecular weight DNA prepared so that there were few nicked strands). Primers had to have a Tm of ~70°C, a GC content of 45-60%, and a low 3' GC content. 1µl of DNA (100ng/µl), 1µl forward primer (10µM), 1µl reverse primer (10µM), 5µl (10x) Advantage2 PCR buffer, 1µl (50x) dNTP mix (10mM each), 1µl (50x) Advantage2 polymerase mix, and 40µl water were added to a 0.2ml microcentrifuge tube. For an expected size product of 1-5kb, the following PCR

program was used: initial denaturation at 95°C for 1min, then 25-35 cycles of denaturation (95°C 30secs), annealing/extension (68°C 3mins), followed by a final polymerase extension at 68°C for 3mins. For products of a larger expected size, the 68°C annealing/extension time was increased to 6mins for 5-9kb products and 12mins for 10-20kb products.

2.1.24 Mouse Genotyping

Ear clips were taken from 3 week old mice that served to mark individual littermates (due to the position of the clips around the circumference of the ear) and provide tissue for genotype analysis. 600µl NaOH (50mM) was added to each ear clip sample, and these were vortexed and boiled for 10mins. Once the samples had cooled, 50µl Tris pH8 (1M) was added. 1µl of boiled tissue lysate was added directly to a 0.2ml microcentrifuge tube containing 22.5µl master mix (1.1x) (AG gene), 6µM of each primer, made up to a total volume of 25µl with water. PCRs were carried out in a TouchGene PCR Cycler (Touche) using the following program: initial denaturation 94°C for 5mins, then 36 cycles of denaturation at 94°C for 1min, annealing at 50-66°C for 1min, extension at 72°C for 1min, followed by a final extension at 72°C for 5mins.

2.1.25 Colony PCR

Individual bacterial colonies were transferred from agar plates into 0.5ml tubes containing 10µl of H_2O using a toothpick. The samples were boiled for 10 mins, cooled on ice, then 1µl of each sample was used in a PCR reaction (section 2.1.20).

2.1.26 Genome Priming System

2µl GPS buffer (10x), 1µl pGPS2.1 donor DNA (0.02µg), 0.08µg target plasmid DNA were placed together in a 1.5ml microcentrifuge tube and made up to a total volume of 18µl with distilled water. 1µl of TnsABC* Transposase was then added and the reagents incubated for 10 mins at 37°C. 1µl of Start Solution was added and a further incubation for 1 hour at 37°C was carried out. The reaction

was stopped by heat inactivation for 10 mins at 75°C. A 10 fold dilution with distilled water was carried out and 1µl and 10µl were used in a transformation reaction with electrocompetent cells. Following transformation, the culture was plated onto LB-agar plates containing chloramphenicol antibiotic and grown overnight at 37°C. The following day, colonies were streaked out onto duplicate LB-agar plates, one containing chloramphenicol antibiotic and the other containing ampicillin antibiotic, and grown overnight at 37°C. Colonies that were able to grow on the chloramphenicol-containing but not the ampicillin-containing LB-agar plates contained the required transposition.

2.1.27 QuikChange® Site Directed Mutagenesis Kit

The QuikChange® Site Directed Mutagenesis Kit was used to introduce specific point mutations into plasmid DNA. Two purified complementary oligonucleotide primers containing a modified nucleotide sequence surrounded by an unmodified nucleotide sequence were designed following manufacturer's instructions. 5µl reaction buffer (10x), 5-50ng dsDNA template, 125ng oligonucleotide primer 1, 125ng oligonucleotide primer 2, 1µl QuikChange® dNTP mix were added to a 0.2ml microcentrifuge tube and made up to 50µl with double distilled water. 1µl *PfuTurbo* DNA polymerase (2.5U) was then added and a PCR reaction was carried out in a TouchGene cycler (Techne) as follows: 95°C for 30 sec, followed by 12 cycles of 95°C for 30 sec, 55°C for 1 min, 68°C for 1 min per 1kb of plasmid length. Reactions were quickly cooled to ≤37°C on ice for 2 mins. 1µl of *DpnI* restriction endonuclease was then added to the PCR reaction, pipetted thoroughly, centrifuged to collect the contents of the tube to the bottom and incubated for 1 hour at 37°C to allow digestion to proceed. 1µl of the *DpnI* treated DNA was used in a transformation reaction with XL1-Blue supercompetent cells following manufacturer's guidelines. Resulting colonies were screened by sequencing across the designed mutagenesis site.

2.1.28 RNA probe synthesis for *in situ* hybridization

A DIG RNA Labelling Kit (Roche) was used to make riboprobes. Plasmid DNA was first linearized with a unique restriction endonuclease positioned at the end of the cloned insert sequence furthest away from the RNA polymerase site to be

used in the RNA probe synthesis reaction. Linearized plasmid DNA was then purified from a 1-1.5% agarose gel using 'Wizard® SV gel clean-up system' (Promega), following the manufacturer's instructions. To make an antisense-specific probe (positive probe), plasmid DNA was digested at the 5' end of the cloned insert sequence. To make a sense-specific probe (negative probe), plasmid DNA was digested at the 3' end of the cloned insert sequence. The following reaction was carried out with RNase-free reagents under RNase-free conditions. Linearized DNA, 2µl of 10x NTP labelling mix (Roche; 10mM ATP, CTP, GTP, 6.5mM UTP, 3.5mM DIG-11-UTP, pH 7.5), 2µl of 10x transcription buffer (Roche), 1µl of 20U/µl RNase inhibitor (Roche) and 2µl of 20U/µl RNA polymerase (Roche) was placed into a 2ml round-bottomed RNase-free tube and incubated at 37°C in a water bath for 2 hours. 2µl of 2U/µl DNase I was added, then a further incubation at 37°C for 15 mins was carried out. 2µl EDTA (0.2M pH8.0) was added to stop the reaction followed by 2.5µl LiCl (4M) and 75µl ice cold 100% ethanol. The tube was incubated at -80°C for 30 mins then centrifuged at 4°C for 15 mins at 12,000rcf. The pellet was washed with 50µl ice cold 70% ethanol, allowed to air dry then resuspended in 19µl nuclease-free water, 1µl EDTA (0.2M pH8.0) and 80µl hybridization buffer (40% formamide, 1x Denhardt's (0.5g ficoll, 0.5g polyvinylpyrrolidone, 0.5g BSA), 5x SSC, 100µg/ml tRNA (Sigma), 100µg/ml herring sperm (Sigma)).

2.1.29 *In situ* hybridization of sectioned tissues

The following procedures were carried out with RNase-free reagents under RNase-free conditions. Slide racks and glass dishes had been baked overnight at 180°C prior to use. Tissue sections of 8µm in thickness on SuperFrost plus coated slides (see section 2.3.3) were placed into a rack and hydrated by submerging into histoclear for 2x 10 mins, 100% ethanol for 2x 5 mins followed by a descending ethanol series (95%, 90%, 80%, 70%, 50%, 30%) of 1 min each. Slides were washed in PBS for 1 min, fixed in 4% PFA for 30 mins, followed by washes in PBS for 2x 1 min. Proteolysis was carried out by incubating the slides in proteinase K (0.025mg/ml) for 15 mins. Slides were washed in PBS for 2x 5 mins, re-fixed in 4% PFA for 15 mins followed by washes in PBS for 2x 5 mins. Slides were then incubated in 2x SSC for 2x 5 mins, Tris-glycine buffer (0.5M

Tris, 0.5M Glycine) for 30 mins, washed in PBS for 30 secs and dehydrated in an ascending ethanol series (30%, 50%, 70%, 80%, 90%, 95%) for 30 secs each followed by 100% ethanol for 2x 1 min and air drying for 30 mins. 500µl hybridization buffer (40% formamide, 1x Denhardt's (0.5g ficoll, 0.5g polyvinylpyrrolidone, 0.5g BSA), 5x SSC, 100µg/ml tRNA (Sigma), 100µg/ml herring sperm (Sigma)) and 2.5-5µl riboprobe (see section 2.1.28), preheated to 95°C for 5 mins and cooled immediately on ice was added to the surface of each slide and overlaid with parafilm ensuring that no bubbles had formed underneath the parafilm. Slides were then incubated overnight at 60°C on a raised platform within an enclosed chamber containing 25ml hybridization chamber buffer (12.5% formamide, 5x SSC) in order to maintain a humid atmosphere. The following procedures were not carried out under RNase-free conditions. Slides were submerged in prewarmed 5x SSC to allow the parafilm to float off the surface of the slides. Slides were then washed with 5x SSC for 3x 15 mins, post-hybridization buffer (20% formamide, 0.5x SSC) for 40 mins at 55°C (5°C lower than the hybridization temperature), 2x SSC for 15 mins, RNase A solution (12.5µg/ml RNase A, 2x SSC) for 15 mins at 37°C, 2x SSC for 15 mins, post-hybridization buffer for 20 mins at 55°C (5°C lower than the hybridization temperature) and 2x SSC for 2x 15 mins. 500µl of 10% blocking buffer (Roche) and 0.25-1µl Anti-DIG AP fab fragments (Roche; 0.75U/µl) prechilled on ice was added to the surface of each slide and overlaid with parafilm ensuring that no bubbles had formed underneath the parafilm. Slides were then incubated overnight at 4°C on a raised platform within an enclosed chamber containing 25ml chamber buffer (2x SSC) in order to maintain a humid atmosphere. Slides were submerged in PBS to allow the parafilm to float off the surface of the slides. Slides were then washed with PBS for 3x 10 mins, 3x 30 mins, 2x 1 hour and NTMT buffer (100mM NaCl, 50mM $MgCl_2$, 100mM Tris-HCl, 0.1% Triton X-100) for 3x 10 mins. 500µl BM purple (Roche) was added to the surface of each slide which were then incubated overnight at room temperature on a raised platform within an enclosed dark chamber containing 25ml milliQ water in order to maintain a humid atmosphere. Slides were checked for the appearance of a staining reaction and BM purple was replaced daily for up to 4 days until this reaction had occurred. The staining reaction was stopped by submerging the

slides in PBS containing 1mM EDTA pH8.0. Finally, slides were dehydrated through an ascending ethanol series for 3 secs in 70% ethanol, 5 secs in 95% ethanol, 2x 1 min in 100% ethanol, washed twice in histoclear clearing solution for 2 mins each and mounted using distyrene, plasticiser and xylene (DPX) mount (Sigma) overlaid with a coverslip.

2.2 Tissue Culture Protocols

(Sambrook et al., 1989)

(Nagy et al., 2003)

2.2.1 Growing ES cells

The feeder-independent, hypoxanthine phosphoribosyltransferase (HPRT)-deficient ES cell line E14Tg2A-clone4 (E14C4) was used as a parental cell line. These cells were derived from the 129P2 (formally 129/OlaHsd) strain of mouse by Dr Austin Smith, University of Edinburgh (ES cells were obtained from Dr William Skarnes and are now available through The Transgenic Core). ES cells were grown without feeder cells in sterile plastic tissue culture dishes (Nunc) coated with 2% gelatin (Life Technologies) with supplemented DMEM growth media: Dulbecco's modified eagle medium (DMEM), with glutamax^{-1}, without sodium pyruvate, with 4500mg/l glucose, with pyridoxine [final conc. 0.08x] (Gibco BRL) supplemented with 15% foetal calf serum (FCS) (Perbio), 1x minimum essential medium (MEM) non-essential amino acids (Gibco BRL), 100μM β-mercaptoethanol, 50μg/ml penicillin/streptomycin (Gibco BRL), 0.1μl/ml ESGRO (10^7units/ml) leukaemia inhibitory factor (LIF) (Chemicon International), and 0.08x sterile milliQ water. The presence of LIF being necessary to prevent differentiation within a feeder-free environment. ES cells were harvested or passaged by washing with PBS followed by the addition of a small amount (enough to cover the bottom of the flask) of trypsin-EDTA (Gibco BRL, 0.05 % Trypsin, 0.53 mM EDTA•4Na), and allowing the cells to become detached from the bottom of the culture dish and disaggregated into a single cell suspension. Growth media was then added to stop the action of trypsin and the cells pelleted by spinning at 180rcf for 4mins. To passage, cells were replated at

a lower density (typically a 1 in 10 dilution) onto a clean gelatin coated tissue culture dish.

2.2.2 Freezing/Thawing ES cells

ES cells were disaggregated with trypsin and pelleted (section 2.2.1), then resuspended in 1ml of 1x freezing medium (supplemented DMEM growth media with 10% DMSO). The cells were then frozen slowly at –80°C in cryovials before plunging into liquid nitrogen. Slow freezing at –80°C was achieved by wrapping cryovials in tissue placed in a polystyrene box. When thawing cells, cryovials were warmed quickly by placing them directly into a 37°C incubator. The DMSO was diluted by adding supplemented DMEM growth media, and cells centrifuged at 180rcf for 4mins to form a pellet. The pellet was then resuspended in 10ml supplemented DMEM growth media and seeded in a gelatin coated T25 (culture area 25cm^2) tissue culture flask.

2.2.3 Freezing/Thawing ES cells in 96-well plates

The protocol as described in Nagy et al. (2003) was followed. Briefly, ES cells were disaggregated with trypsin into a single cell suspension (section 2.2.1). The action of the trypsin was then stopped by adding a small volume of supplemented DMEM growth media. The cells were transferred into a U-bottomed 96-well plate (Nunc) placed on ice and an equal volume of 2x freezing medium (supplemented DMEM growth media with 20% DMSO) was added to the cells, which was then overlaid with mineral oil (Sigma). The plates were wrapped in Parafilm and placed in a polystyrene box at -80°C.

2.2.4 Rapid preparation of DNA from cells in 96-well plates and 24-well plates

The protocol as described in Nagy et al. (2003) was followed for preparing DNA from 96-well plates (with DNA resuspended in 25µl TE before restriction digestion). For DNA preparation from 24-well plates, the same protocol was followed as for preparing DNA from 96-well plates, but all volumes were increased 4-fold (with DNA resuspended in 150µl TE before restriction

digestion). Briefly, ES cells were lysed directly in tissue culture plates in the presence of a lysis buffer containing proteinase K. The dishes were incubated overnight at 60°C in a humid environment in a rocking incubator and the following day a mixture of NaCl and cold ethanol was added to precipitate the DNA. The precipitated DNA was then washed in 70% ethanol before being resuspended in TE.

2.2.5 Selective antibiotic dose response

The dose of Geneticin® G418 sulphate (Fluorochem Ltd.) used for transfection of the parental ES cell line (E14C4) was determined by an antibiotic dose response test. G418 was added to supplemented DMEM growth media to a concentration of 5mg/ml and sterilized by filtering through a 0.22 micron filter. ES cells were grown in a large T75 (culture area 75cm^2) gelatin-coated tissue culture dish until ~80% confluent. The cells were harvested and the cell density counted using a haemocytometer (section 2.2.11). Approximately 400 cells were then added to each of eleven gelatin-coated tissue culture petri dishes (35mm diameter) containing supplemented DMEM growth media. The following day G418 was added to each dish of cells at concentrations ranging from 0-1000μg/ml. The G418 containing medium was changed every 4-5days, or more often if there had been a high rate of cell death (indicated by cell debris visible in the culture medium). After 10-14 days, the cells were washed with PBS and the concentration of G418 that gave almost total cell death was determined. A second dose response test was then carried out over a smaller concentration range, in increments of 25μg/ml, in order to home in on the optimum G418 concentration to use in order to kill all non-antibiotic resistant cells. After 10-14 days the G418 concentration that gave almost complete cell death was again determined. To aid visualization of any remaining living cells, cells were fixed in 10% formaldehyde for 1 hour followed by overnight staining with 33% methylene blue. The optimum G418 concentration used to kill ~100% of non-transfected E14C4 ES cells was determined to be 225μg/ml. This concentration of G418 was used in transfection experiments.

2.2.6 ES cell transfections

E14C4 cells were grown in T175 (culture area 175cm^2) tissue culture flasks until ~80% confluent. Growth medium was changed 2-3 hours before harvesting cells. The cells were disaggregated using trypsin and pelleted (section 2.2.1). The pellet was then washed with PBS and resuspended in a small volume of Ca++/Mg++ -free PBS. The cell density was counted using a haemocytometer (section 2.2.11) and the cell suspension adjusted to a density of ~1 x 10^7 cells/ml. 0.8ml of the single-cell suspension was placed into a 0.4cm-wide electroporation cuvette (Bio-Rad). 20-25µg of linearised targeting construct DNA (an amount recommended by Nagy et al. (2003)) purified using the 'Wizard® SV gel clean-up system' (Promega) was added to the cells and allowed to stand at room temperature for 5mins. A single electrical pulse (300V, 250µF) was applied to the cuvette using a GenePulser electroporator (Bio-Rad). The cells were allowed to stand for a further 5mins at room temperature. The electroporated cells were divided between twelve gelatin-coated tissue culture petri dishes (10cm diameter) containing supplemented DMEM growth media and incubated at 37°C, 5%CO_2. The following day the growth medium was changed and G418 was added to a concentration of 225µg/ml to all but one of the dishes (this final dish acted as a control for antibiotic selection). The G418-containing medium was changed every 4-5 days or more often if there was a lot of cell death.

After 11-12 days any remaining live cell colonies in the G418 containing dishes should have incorporated the targeting construct, and therefore acquired geneticin resistance. Individual colonies were picked as follows: The growth medium was aspirated and the cells rinsed with Ca++/Mg++ -containing PBS, leaving just enough to cover the surface of the dish. Colonies were picked individually by dispensing 10µl of trypsin-EDTA (Gibco BRL, 0.05 % Trypsin, 0.53 mM EDTA•4Na) briefly on top of each cell clump (cell clumps were visible by eye after 10-12 days in culture), then transferring any detached cells to a single well of a 96-well U-bottomed uncoated tissue culture dish containing 20µl trypsin-EDTA. After 15mins the 30µl of trypsin and cells were transferred to a single well of a 96-well flat-bottomed gelatin coated tissue culture dish containing 100µl supplemented DMEM plus G418 (225µg/ml). The medium was

changed each day. After 3-4 days the cells had grown sufficiently for passaging, and expanding. Duplicate 96-well plates were grown until there were enough cells to freeze one plate for storage, make DNA from a second plate and expand cell numbers from a third plate.

2.2.7 Genotyping transfected ES cell colonies

Transfected ES cells were genotyped either by Southern hybridization or by PCR. For genotyping by Southern hybridization, cells were grown in 24-well plates until 80-90% confluent and DNA was prepared (section 2.2.4). DNA was then digested directly in each well of the 24-well plates. The following reagents were added to the 150µl of TE in each well: 18µl (10x) restriction enzyme buffer (Promega), 1.8µl (100x) BSA (Promega), 6µl (10-12units/µl) restriction enzyme (Promega) and 4.2µl water. The DNA was then incubated overnight at 37°C (or at the appropriate temperature for the restriction enzyme), in a rocking incubator. 150-200µl of each digested DNA sample was added to individual wells of a 1% agarose gel and electrophoresed overnight at 25V. The agarose gel was blotted as described in section 2.1.12. The Southern blot was then hybridized with a radiolabelled probe that recognises a restriction length polymorphism between the endogenous and targeted alleles if the targeted construct had incorporated into the correct genomic location. For genotyping by PCR, DNA was prepared from 96-well plates (section 2.2.4) and 1µl was used for each PCR reaction (section 2.1.20).

2.2.8 LacZ staining of ES cells

ES cells were grown on gelatinised petri dishes (35mm diameter) without feeder cells (section 2.2.1) until ~80% confluent. Cells were washed twice with PBS, 2mM $MgCl_2$ for 5mins then fixed for 5mins in 0.8% glutaraldehyde followed by three 5min washes with PBS, 2mM $MgCl_2$. Cells were then treated with permeabilisation solution (0.1% Triton X-100, 2mM $MgCl_2$, in PBS) for 15mins followed by a further 5min wash step with PBS, 2mM $MgCl_2$. Finally, cells were incubated overnight at 28°C in X-gal stain base reagent (1mg/ml X-gal in N,N-Dimethylformamide, 2mM ($MgCl_2$), 5mM potassium ferricyanide ($K_3Fe(CN)_6$),

5mM potassium ferrocyanide ($K_4Fe(CN)_6$), 0.24mM sodium deoxycholate ($C_{24}H_{39}NaO_4$), in PBS, pH7.4-7.5).

2.2.9 Preparation of chromosome spreads

ES cells were grown on gelatinised petri dishes (35mm diameter) without feeder cells (section 2.2.1) until ~80% confluent. Cells were then passaged 1:2 into another 35mm dish one day prior to performing the chromosome spreads to ensure actively dividing cells were present. Cells were medium changed 2-3 hours prior to the addition of Colcemid (GibcoBRL), at a concentration of 0.1μg/ml, and the cells incubated for a further hour at 37°C. The medium was aspirated and the cells rinsed with PBS. The cells were then thoroughly dispersed following the addition of trypsin-EDTA (Gibco BRL, 0.05 % Trypsin, 0.53 mM EDTA•4Na) for 10 to 15mins, to ensure a single cell suspension. An equal volume of media was added to stop the action of the trypsin. Cells were then centrifuged for 10 minutes at 145rcf. The supernatant was poured off and the pellet resuspended in any remaining liquid. 5 ml of 0.56% (0.075M) KC1 was added and left to stand at room temperature for 10 minutes which causes the cells to swell. The cells were spun down for 5 minutes at 145rcf, the supernatant poured off and the cell pellet resuspended thoroughly in the remaining liquid. Using freshly prepared fixative, methanol:glacial acetic acid (3:1) cooled to 4°C, 5 ml was slowly added to the cell suspension (the first 2 ml was added drop by drop, tapping the tube to mix after each addition; the remaining 3 ml was added at a faster rate). The cells were centrifuged for 5mins at 145rcf and the fixation step repeated. The cells were centrifuged for a final time for 5min at 145rcf. The supernatant was removed and the pellet resuspended in ~250μl of fixative. The cells were kept on ice from this point. One drop of cell suspension in fixative was allowed to fall onto a clean slide, from about 10cm above. The slide was left to air dry for 1-2mins then viewed under a microscope to determine if there were too few or too many cells per slide. The cells should be fairly well separated, with clumps of only 2 or 3 cells. If there were too few cells the suspension was spun down and resuspend in a smaller volume; if too many cells, a small amount of fixative was added to the cell suspension. One drop of mounting medium with DAPI (H-1200, Vector Laboratories, Inc., Burlingame, CA) was added to the

spread, and overlaid with a coverslip. The chromosome spreads were viewed using a fluorescent microscope with a 100x oil immersion lens objective (x1000 total magnification). Photographs of metaphase spread were taken with a 35mm camera. The slides were kept for up to 14 days at 4°C in the dark.

2.2.10 RNA extraction from ES cells

An 'RNeasy kit' (Qiagen) was used to extract RNA from ES cell. Cells were grown in a large tissue culture petri dish (140mm diameter) until ~80% confluent. 600µl of 'RLT buffer' supplemented with 6µl β-mercaptoethanol was added to lyse the cells. The cells were detached from the bottom of the dish using a cell scraper (Corning Incorporated) and transferred to an RNase-free 1.5ml microcentrifuge tube, then vortexed until no visible cell clumps remained. The cells were homogenized by pipetting the lysate into a 'QIAshredder spin column' (Qiagen) and centrifuging for 2mins at 10,000rcf. 1 volume (~700µl) of 70% ethanol was added to the homogenized cells. 200µl of the sample was then added to an 'RNeasy mini column' and centrifuge for 15 seconds at 8000rcf (if a large quantity of RNA was required, more than one column was used). The flow-through was discarded and 700µl of 'RW1 buffer' was added. The column was centrifuged for a further 15 seconds at 8000rcf, then transferred to a fresh collection tube. 500µl 'RPE buffer' was added, then the column centrifuged for 15 seconds at 8000rcf. The flow-through was discarded, and a further 500µl 'RPE buffer' was added. The column was centrifuged for 2mins at 8000rcf, the flow-through discarded, then spun dry for 1min at 8000rcf . The column was then transferred to a clean RNase-free 1.5ml microcentrifuge tube. To elute the RNA from the column, 30µl sterile RNase-free water was added and the column centrifuged for 1min at 8000rcf. The elution step was then repeated with a further 30µl of water. RNA was stored at –80°C. To get rid of any contaminating DNA, the 'DNA-*free* kit' (Ambion) was used. 6.7µl of (10x) DNaseI buffer and 2µl DNaseI was added to the RNA and incubated at 37°C for 30mins. The DNaseI was inactivated by adding 7µl DNase inactivation reagent and incubating at room temperature for 2mins. The RNA was then centrifuged for 1min at 8000rcf to pellet the inactivation reagent and the supernatant transferred to a fresh tube.

2.2.11 Counting ES cells using a haemocytometer

The haemocytometer was cleaned with 70% ethanol and a lightly moistened coverslip was pressed firmly over the haemocytometer chamber. ES cells were harvested as described in section 2.2.1 and the pellet resuspended thoroughly in a small volume of supplemented DMEM growth media. A small volume of cell suspension (enough to cover the polished surfaces, ~10µl) was pipetted into each of the two haemocytometer chambers. The distribution of cells was observed under 200x magnification to check that they were not too sparse or dense. If necessary, the volume of cell suspension was adjusted accordingly or a dilution of the cell suspension was placed onto the haemocytometer. A hand held counter was used to count the number of cells located in the central square of each chamber (5x5 square surrounded by triple lines) and also the corner four squares of each chamber (4x4 square) if a more accurate count was needed. To prevent counting each cell more than once, cells were only counted if they lay within the square or were touching the top and left of the middle line. The average number of cells within these squares was calculated; this was the number of cells per 10^{-4}ml. The total number of cells per ml was therefore = cells per 10^{-4}ml x $1x10^4$ x dilution factor (the dilution factor being = 1 if no dilution of the cell suspension was carried out). The total number of cells in the original suspension was therefore calculated to be = total number of cells per ml x total volume.

2.3 Histology Protocols

2.3.1 Fixation and wax embedding of tissues

Tissues were dissected and placed into 4% PFA overnight at 4°C. The following day organs were rinsed briefly in PBS then transferred to 70% ethanol until ready to be processed (no later than 72 hours following fixation). Processing was carried out in a Leica TP1020 machine with 2x 1 hour washes in 70% ethanol, 90% ethanol, 100% ethanol, histoclear clearing solution, followed by 3x 1 hour washes in 40°C molten wax (Raymond Lamb). Following processing, tissues were embedded in 40°C molten wax within histology cassettes and allowed to cool until solid on a 4°C cooled platform embedding station (Leica, EG1160).

2.3.2 Subbing slides with APTS

Slides were initially washed in hot soapy water followed by thorough rinsing in hot water and a quick rinse in reverse osmosis (RO) water then in a solution of 95% ethanol, 01% acetic acid and allowed to dry at room temperature. Slides were then placed into 2% 3-triethoxysilylpropylamine (APTS) in acetone for 10 mins followed by rinsing twice in acetone for 10 secs then twice in RO water for 10 secs. Finally slides were baked dry at 37°C.

2.3.3 Sectioning of wax embedded tissues

Sections of 8μm in thickness were cut using a microtome (Leica, RM2155) and placed onto SuperFrost plus coated slides (Fisher) (for *in situ* hybridization) or APTS subbed slides (for histological staining), dried at 38°C overnight then stored at 4°C. All sectioning was performed by Iryna Withington, University of Bath.

2.3.4 Haematoxylin and eosin (H&E) staining of sectioned tissues

Slides of 8μm wax embedded organ sections were washed twice in histoclear clearing solution (National Diagnostics) for 2 mins to remove wax. The slides were then hydrated through a descending ethanol series for 2x 1 min each in 100% ethanol followed by 1x 1 min each in 95% ethanol, 90% ethanol, 70% ethanol, 50% ethanol. The slides were placed in milliQ water for 1 min before staining in haematoxylin for 15 mins. To eliminate non-specific binding of haematoxylin, slides were placed under gently running tap water for 3 mins, followed by a 30 sec wash in 1% concentrated hydrochloric acid (HCl) in 70% ethanol, a 1 min wash in 1% ammonia (NH_3) in 70% ethanol and a final 30 sec wash in 70% ethanol. Slides were then counterstained in a solution of filtered eosin for 5 mins. Finally, slides were dehydrated through an ascending ethanol series for 3 secs in 70% ethanol, 5 secs in 95% ethanol, 2x 1 min in 100% ethanol, washed twice in histoclear clearing solution for 2 mins each and mounted using DPX mount (Sigma) overlaid with a coverslip.

2.3.5 Periodic acid-Schiff (PAS) staining

Slides with 8μm wax embedded organ sections were washed twice in histoclear clearing solution (National Diagnostics) for 2 mins to remove wax. The slides were then hydrated through a descending ethanol series for 2x 1 min in 100% ethanol followed by 1x 1 min each in 95% ethanol, 90% ethanol, 70% ethanol, 50% ethanol. The slides were placed in milliQ water for 1 min then transferred to freshly prepared 1% periodic acid (BHD) for 6 mins. Slides were washed by placing them under gently running tap water for 3 mins then into milliQ water for 1 min. Staining was carried out at room temperature in Schiff's reagent (Sigma) for 15 mins. Slides were washed by placing them under gently running tap water for 5 mins then counterstaining in haematoxylin for 1.5 mins. Finally slides were dehydrated through an ascending ethanol series for 1x 1 min each in 70% ethanol, 95% ethanol, 2x 1 min in 100% ethanol, washed twice in histoclear clearing solution for 2 mins each and mounted using DPX mount (Sigma) overlaid with a coverslip.

2.3.6 Decalcification of bone

Bone tissue was placed in decalcification solution (22% formic acid, 10% sodium citrate dihydrate ($Na_3C_6O_7H_5 \cdot 2H_2O$)) at 4°C overnight or longer. Decalcification was judged to be complete when a high gauge needle could easily pierce the bone. The tissue was washed under running tap water for several minutes before being transferred into 70% ethanol then processed by fixation and wax embedding.

2.4 Mouse Protocols

(Sambrook et al., 1989)

(Nagy et al., 2003)

2.4.1 Total RNA extraction from mouse tissue

The appropriate tissues were excised from a mouse. An 'RNeasy kit' (Qiagen) was used to extract RNA from mouse tissue. The tissues were weighed and if over 30mg, several 25mg sections were processed separately. The tissue was

placed into an RNase-free tube and 600µl 'RLT buffer' supplemented with 6µl β-mercaptoethanol was added. The sample was homogenized with a hand held rotor-stator homogenizer (Jencons), until there were no remaining clumps of tissue (the homogenizer probes were cleaned before use with 3% hydrogen peroxide to render them RNase-free). The lysates were then centrifuged for 3mins at 10,000rcf and the supernatant transferred into a fresh tube. 1 volume of 70% ethanol was added to the supernatant and mixed. 700µl of each sample was then added to an 'RNeasy mini column' and centrifuged for 15 seconds at 8000rcf. The flow-through was discarded and 700µl 'RW1 buffer' was added to the columns then centrifuged for 15 seconds at 8000rcf. The columns were then transferred to fresh collection tubes. 500µl 'RPE buffer' was added to each column, centrifuge for 15 seconds at 8000rcf and the flow-though discarded. A further 500µl 'RPE buffer' was added and centrifuged for 2mins at 8000rcf. The column was then spun dry for 1min at 8000rcf and transferred to a clean RNase-free 1.5ml microcentrifuge tube. To elute the RNA from the column, 30-50µl sterile RNase-free water was added to the column and centrifuged for 1min at 8000rcf. This elution step was then repeated with a further 30-50µl of water. RNA was stored at -80°C. To remove any contaminating DNA, the 'DNA-*free* kit' (Ambion) was used following the manufacturers instructions.

2.4.2 Poly(A)$^+$ RNA extraction from mouse tissue

The appropriate tissues were excised from a mouse. A 'Micro-FastTrack™ 2.0 RNA kit' (Invitrogen) was used to extract mRNA from mouse tissue. The tissues were weighed and 10-200mg sections were placed into RNase-free 1.5ml microcentrifuge tubes and processed separately. 1ml of Micro-FastTrack™ 2.0 lysis buffer containing proteinase K, preheated to 45°C, was added to each tube and the tissues were homogenized by passing through a 18-21 gauge needle or by using a hand held rotor-stator homogenizer (Jencons), until there were no remaining clumps of tissue (the homogenizer probes were cleaned before use with 3% hydrogen peroxide to render them RNase-free). The cell lysates were then incubated for 20 mins at 45°C and any remaining cell clumps were removed. 63µl of NaCl (5M) was then added to each sample and mixed thoroughly. To remove any remaining DNA the samples were passed through a 18-21 gauge

needle. The samples were added to vials of oligo(dT) cellulose which were subsequently sealed and incubated in an upright position for 2 mins to allow the oligo(dT) cellulose to swell, then transferred to a rocking platform for 20 mins to allow the mRNA to bind to the oligo(dT) cellulose. The vials were then centrifuged for 5 mins at 4000rcf and the supernatant removed. The oligo(dT) cellulose was resuspended in 1.3ml of binding buffer, centrifuged for 5 mins at 4000rcf and the supernatant removed; this step was repeated until the supernatant appeared transparent. The oligo(dT) cellulose was then resuspended in 300µl of binding buffer and transferred to a spin column and centrifuged for 10 secs at 4000rcf and the flow through discarded; this step was repeated until all the oligo(dT) cellulose had been transferred to the spin column. 500µl of binding buffer was then added to the spin column, centrifuged for 10 secs at 4000rcf, and the process repeated until the OD_{260} of the flow through was <0.05. 200µl of low salt wash buffer was added to the spin column and the oligo(dT) cellulose carefully resuspended with a pipette then centrifuged for 10 secs at 4000rcf and the flow through discarded; this step was repeated. The spin column was placed into a fresh 1.5ml microcentifuge tube. 100µl of elution buffer was added to the spin column and the oligo(dT) cellulose carefully resuspended with a pipette then centrifuged for 10 secs at 4000rcf; this step was repeated. The mRNA was now in the flow through. This mRNA was then precipitated on ice for 10 mins by adding 10µl glycogen carrier (2mg/ml), 30µl sodium acetate (2M) and 600µl ethanol (100%). The sample was centrifuged at 4°C for 15 mins at 10,000rcf, the supernatant discarded, the pellet washed in 70% ethanol followed by air drying for 5-10 mins, and finally resuspended in 10µl of elution buffer. 1µl of cDNA was used for each RT-PCR reaction.

2.4.3 ES cell injections into mouse blastocysts

The schedule for producing chimeric mice by injection of mouse blastocysts with ES cells is summarised in Table 2.1. It was found after much practice and experimentation that blastocysts obtained from naturally mated C57BL/6 mice were optimum for injection, their cell membranes being much easier to pierce with the injection needle. Other parameters were also tried and altered until the optimum conditions, as outlined below, were found.

Table 2.1

Summary of the schedule for producing chimeric mice by injection of mouse blastocysts with ES cells

Day 1	Administration of PMSG to blastocyst donors (superovulated females)
Day 2	
Day 3	Administration of hCG to blastocyst donors (superovulated females) Set up blastocyst donors for mating with stud males (superovulated and naturally mated females) Thaw ES cells and grow in culture
Day 4	Check for copulation plugs in blastocyst donor females Set up pseudopregnant recipients for mating with vasectomized males ES cells were disaggregated with trypsin and seeded at a lower density
Day 5	Check for copulation plugs in pseudopregnant recipients
Day 6	
Day 7	Harvest blastocysts Enrich and harvest ES cells Blastocyst injections and transplantation into pseudopregnant recipients
Day 8	Check for wound closure in pseudopregnant recipients
Day 22	Check for signs of pregnancy (15 days post-transfer)
Day 24	Check for pups born (typically 17 days post-transfer)

F2 (C57Bl/6 X CBA) or C57Bl/6 mice were used as blastocyst donors. For each round of blastocyst injections either 5-7 female mice (3.5-4.5 weeks of age) were superovulated before mating, or ~20 female mice (6 weeks - 4 months of age) were natural mated in order to provide blastocyst embryos for injection. 1-2 females were placed in cage overnight with a single stud male and checked for the presence of a copulation plug the next day. Typically, 4-5 mice out of the 20 set up for natural matings showed presence of a copulation plug the following day; the mouse oestrus cycle is 4 days in length and therefore, not all the females were in oestrus at the same time. For superovulated mice, hormone gonadotrophins were administered via intraperitoneal (IP) injections to coincide with the light-dark cycle (12/12 cycle; 7am-7pm). Female mice were injected 46-48 hours prior to ovulation (at ~2.30pm on day 1) with 5IU PMSG, and 10-13 hours prior to ovulation (at ~12.30 on day 3) with 5IU of hCG (Nagy et al., 2003). Mice were set up for mating in the afternoon following administration of hCG (day 3) and were assumed to have mated at the mid-point of the dark cycle. Hormones were made up as follows: PMSG stock solutions were made by reconstituting a 5000IU vial of PMSG (Intervet UK) in 10ml sterile water. 1ml (500IU) aliquots were then placed into ten 1.5ml microcentrifuge tubes and frozen at –20°C for up to 6 months. When fresh working stock solutions were required, a 500IU aliquot of PMSG was thawed and 100µl aliquots were placed into ten 1.5ml microcentrifuge tubes, 900µl of sterile water was added to each tube to make a 50IUstock. When required, a 50IU tube was thawed and 0.1ml (5IU) injected per mouse. hCG working stock solutions were made by reconstituting a 1500IU vial of hCG Chorulon® (Intervet UK) in 750µl sterile water. Aliquots of 25µl were placed into 1.5ml microcentrifuge tubes and dried for 45 minutes in a centrifuge under vacuum (Speedvac Concentrator, Stratech) then stored at –20°C. When required, a single pellet was re-dissolved in 1ml sterile PBS, mixed thoroughly, and 0.1ml (5IU) injected per mouse.

3.5days after mating (day 7), blastocysts were collected from the donor females (killed by cervical dislocation) by dissecting the uterine horns from the body cavities and flushing out the embryos using a 30 gauge needle attached to a 1ml syringe filled with FHM medium. FHM medium: 5ml stock A' (950mM NaCl,

25mM KCl, 3.5mM KH_2PO_4, 2mM $MgSO_4$, 2mM glucose, 100mM sodium lactate, 600µg/ml penicillin, 500µg/ml streptomycin), 0.8ml stock B (250mM $NaHCO_3$, 0.001% phenol red), 0.5ml stock C' (20mM sodium pyruvate), 0.5ml stock D (17.1mM $CaCl_2.2H_2O$), 4ml stock E (250mM HEPES, 0.001% phenol red), 5µl stock F (100mM $Na_2EDTA.2H_2O$), 250µl stock G (200mM glutamine), 39ml milliQ water, 200mg BSA, pH adjusted to 7.2-7.4 with 390µl sodium hydroxide, filter sterilised through a 0.45µm filter (Millipore). Expelled blastocysts were collected using a transfer pipette attached to a mouth pipette assembly (Sigma) and transferred into a drop of KSOM medium equilibrated with CO_2, overlaid with mineral oil (Sigma), then incubated at 37°C, 5% CO_2. KSOM medium (stocks the same as for FHM medium): 1ml stock A', 1ml stock B, 100µl stock C', 100µl stock D, 1µl stock F, 50µl stock G, 7.8ml milliQ water, 10mg BSA, equilibrated with CO_2 to obtain correct pH, filter sterilised through a 0.22µm filter (Millipore). Transfer pipettes were made in the laboratory using borosilicate glass capillaries (1mm outer diameter x 0.78mm inner diameter, without filament; Harvard apparatus). They were made by pulling the glass capillaries over a standard bunsen flame then bending the stretched glass until it snaps; only cleanly snapped pipettes were used as jagged edges could damage the blastocysts.

ES cells to be injected (originally derived from a 129P2 stain of mouse) were grown until 60-80% confluent, then enriched for undifferentiated cells 1.5-2 hours before required. Enrichment was achieved by tyrpsinizing the cells and replated them at a lower density (typically a 1:4 dilution) into a clean gelatinized tissue culture flask. After this incubation time, any remaining floating cells that hadn't attached themselves to the bottom of the dish were collected and pelleted by centrifugation for 4mins at 180rcf (cells that were quick to attach were assumed to be mainly differentiated cells). The pellet was resuspended in 1ml KSOM medium and centrifuged for a further 4mins at 180rcf. The supernatant was discarded and the pellet resuspended in 100µl KSOM medium and kept either at room temperature or on ice.

Injections were carried out using a Nikon Eclipse TE300 inverted microscope, Nikon Narishige motorised joysticks and Nikon IM-16 Microinjectors. A water cooled metal stage was constructed to keep the blastocysts rigid whilst being injected. An injection chamber (lid of a 35mm diameter petri dish) containing a small drop of FHM overlaid with mineral oil was placed onto the cooled stage and a holding pipette and an injection needle were lowered into the medium. Blastocysts were injected using commercially available injection needles (purchased form BioMedical Instruments, number 37, spike 55). Blastocysts were held in place using holding pipettes made in the laboratory using borosilicate glass capillaries pulled over a small bunsen flame, snapping in two, then scored carefully with a diamond-tipped pen, so that the outer diameter was 1.2-2.0units (visualized with 10x magnification under a dissecting microscope). The blunted capillaries were then polished with a heated filament using a Beaudoin Microforge 620, until the internal diameter was 0.1units (visualized with 10x magnification). A slight bend was added close to the end using the heated filament, in order that the holding pipette would sit flat on the base of the injection chamber. ES cells and blastocysts were transferred into the injection chamber using a transfer pipette. Injected blastocysts were transferred to a fresh drop of KSOM and incubated at 37°C, 5% CO_2, until transfer into a pseudopregnant recipient female.

Pseudopregnant recipients (MF1 female mice 6-8 weeks of age) were set up for mating with vasectomized males 2.5 days before transfer of blastocysts (day 4). Females likely to be in oestrus (as judged by vaginal appearance) were placed in a cage overnight with a single vasectomized male, and checked for the presence of a copulation plug the next day. Females were weighed and anaesthetic was administered. Anaesthetic used was a mix of Hypnorm™/Hypovel®. A working concentration of hypnorm/hypnovel was made by combining 2ml Hypnorm™ (0.315mg/ml fentanyl citrate, 10mg/ml fluanisone), 4ml water and 2ml hypnovel® (5mg/ml midazolam) followed by filter sterilisation. A dose of 0.006ml/g (0.18ml per 30g mouse) was administered IP, followed by a 0.2ml dose of Vetergesic® (0.3mg/ml buprenorphine) analgesia administered by subcutaneous (SC) injection after surgery. The anaesthetised mouse was laid on

its side and a small incision of the skin and body wall was made in the mid dorsal region. The ovary, oviduct and uterus were carefully lifted out of the body by pulling on the ovarian fat pad which was held in place outside of the body wall with a crocodile clip. A hole was made into the uterus, just below the utero-tubal junction using a 30 gauge needle. A uterine transfer pipette containing the injected blastocysts and attached to a mouth pipette assembly was then inserted into this hole and the embryos expelled gently. The uterus was placed back into the body cavity and the wound closed using coated Vicryl® braided absorbable suture material with a 16mm curved needle. Uterine transfer pipettes were made in the laboratory using borosilicate glass capillaries pulled over a small bunsen flame, snapped, then polished smooth with a heated filament using a Beaudoin Microforge 620.

2.4.4 ES cell aggregations with mouse morula-stage embryos

The schedule for producing chimeric mice by aggregation of mouse morullae with ES cells is summarised in Table 2.2.

F2 (C57Bl/6 X CBA) mice were used as embryo donors. For each round of aggregations, 5-6 female mice (3.5-4.5 weeks of age) were superovulated before mating, in order to provide morula-stage embryos for aggregations with ES cells. 1-2 females were placed in cage overnight with a single stud male and checked for the presence of a copulation plug the next day. Hormone gonadotrophins were administered as described in section 2.4.3. 2.5 days after mating (day 6), embryos were collected from the donor females by dissecting the uterine horns from the body cavities and flushing out the embryos using a 30 gauge needle attached to a 1ml syringe filled with FHM medium. Expelled morulae embryos were collected using a transfer pipette attached to a mouth pipette assembly (Sigma) and transferred into a drop of KSOM medium equilibrated with CO_2, overlaid with mineral oil, then incubated at 37°C, 5% CO_2.

The zonae of the embryos were removed by placing in acidic Tyrode's solution (Sigma). The embryos were placed in a drop of FHM, washed briefly in a drop of

Table 2.2

Summary of the schedule for producing chimeric mice by aggregation of mouse morullae with ES cells

Day 1	Administration of PMSG to embryo donors (superovulated females)
Day 2	
Day 3	Administration of hCG to embryo donors (superovulated females) + set up for mating with stud males Thaw ES cells and grow in culture
Day 4	Check for copulation plugs in embryo donor females Set up pseudopregnant recipients for mating with vasectomized males ES cells were disaggregated with trypsin and seeded at a lower density
Day 5	Check for copulation plugs in pseudopregnant recipients
Day 6	Harvest embryos Enrich and harvest ES cells Aggregations
Day 7	Transplantation into pseudopregnant recipients
Day 8	Check for wound closure in pseudopregnant recipients
Day 22	Check for signs of pregnancy (15 days post-transfer)
Day 25	Check for pups born (typically 18 days post-transfer)

acidic Tyrode's solution, then transferred to a second drop of acidic Tyrode's solution until the zonae had dissolved. The dish was then immediately flooded with FHM to stop the action of the acidic Tyrode's before transferring the zona-free embryos into a fresh drop of KSOM medium overlaid with mineral oil.

Embryos were transferred into individual wells of an aggregation plate. Aggregation plates were made using 35mm diameter petri dishes containing several drops of KSOM overlaid with mineral oil. Indented wells were made at the bottom of each of these KSOM drops using a metal darning needle. Embryos were transferred into each of these KSOM drops using a transfer pipette, and moved into individual wells using a sweeping needle. Sweeping needles were made by pulling borosilicate glass capillaries into a long, sharp tapered needle using a Sutter Micropipette Puller.

ES cells were added to the KSOM drops to an optimal cell density (determined empirically). ES cells to be used (from a 129P2 stain of mouse) were grown until 60-80% confluent as described in section 2.2.1, then enriched for undifferentiated cells 1.5-2 hours before required, and pelleted in 100µl KSOM medium. Morulae embryos and ES cells were then placed in an incubator at 37°C, 5% CO_2 overnight to allow the ES cells to incorporate with the embryos.

The following day, aggregated embryos were transferred into pseudopregnant recipients which were set up for mating with vasectomized males 2.5 days before transfer (day 4). Mice were anaesthetised and surgery was carried out as described in section 2.4.3.

2.4.5 Breeding Chimeras

Chimeric mice were placed in a cage with a 129 mouse of the opposite sex. A week following birth, the coat colour of pups was observed. If some of the injected ES cells have contributed to the gonads of the chimeric mice, some of the gametes will be of the ES cell-derived genotype. Pups with a chinchilla coat

colour arise from gametes of the ES cell-derived genotype, as a chinchilla coat colour is a recessive trait. An ear clip was taken from chinchilla mice at three weeks of age in order for them to be genotyped to ascertain whether the ES-cell derived genotype was wildtype or mutant. Pups with an agouti coat colour arise from gametes of the host chimeric genotype and were therefore culled at three weeks of age.

2.4.6 *In vitro* fertilization

Male mice were placed alone in a clean cage three to six days prior to being used as sperm donors. Female F1 hybrid mice (3.5-4.5 weeks of age) to be used as oocyte donors were superovulated by the administration of 5IU PMSG via IP injection, followed by the administration of 5IU hCG via IP injection 48 hours later. Coincidence of hormone injections with the light-dark cycle was not necessary.

The day prior to ooctye collection, all fertilization and culture dishes were prepared, overlaid with mineral oil and incubated at 37°C, 5% CO_2 in order for the medium to be sufficiently heated and pre-gassed. A male mouse (either F1 hybrid or chimeric, >8 weeks of age) was sacrificed 12-13 hours after the females were hCG injected, the cauda epididymis and vas deferens were removed and placed into 1ml of human tubal fluid (HTF) (101.6mM NaCl, 4.69mM KCl, 0.37mM KH_2PO_4, 0.2mM $MgSO_4.7H_2O$, 21.4mM Na lactate, 0.33mM Na pyruvate, 2.78mM glucose, 25mM $NaHCO_3$, 2.04mM $CaCl_2.2H_2O$, 0.075mg/ml penicillin-G, 0.05mg/ml streptomycin sulfate, 0.4ml 0.5% phenol red solution, 4mg/ml BSA) overlaid with mineral oil. Several slits were made in the cauda epididymis with a 30G needle, and the vas deferens was squeezed to release the sperm. The dish was incubated for 10 minutes at 37°C, 5% CO_2, followed by transferring 10µl aliquots of sperm to 500µl drops of HTF overlaid with mineral oil using a wide bore pipette tip. 13 hours after hCG injections, females were sacrificed, the ovaries and oviducts were transferred to a dish containing PBS. Swollen apullas were torn and the clutches of oocytes held together by the cumulus cells were transferred in a minimum amount of PBS to

the 500µl HTF already containing the sperm. These IVF dishes, were incubated at 37°C, 5% CO_2 for 4-6 hours, with the lid of the dishes removed to guard against a seal of oil being formed thereby preventing gas equilibration from being maintained. After this incubation period all potentially fertilized oocytes were washed in a series of 150µl KSOM-AA (1ml stock A' (950mM NaCl, 25mM KCl, 3.5mM KH_2PO_4, 2mM $MgSO_4$, 55.6mM glucose, 100mM sodium lactate, 600µg/ml penicillin, 500µg/ml streptomycin), 1ml stock B (250mM $NaHCO_3$, 0.001% phenol red), 100µl stock C' (20mM sodium pyruvate), 100µl stock D (17.1mM $CaCl_2.2H_2O$), 1µl stock F (100mM $Na_2EDTA.2H_2O$), 50µl stock G (200mM glutamine), 50µl 100x MEM non-essential amino acids (Gibco BRL), 100µl 50x MEM amino acids (Gibco BRL), 7.8ml milliQ water, 40mg BSA, equilibrated with CO_2 to obtain correct pH, filter sterilised through a 0.22µm filter (Millipore)) drops until all debris and sperm were removed then incubated overnight at 37°C, 5% CO_2.

The following day, two-cell stage embryos were surgically re-implanted into the oviducts of pseudopregnant MF1 recipients that had been mated with vasectomized males the previous evening (section 2.4.3 for details of selecting and anesthetizing pseudopregnant females and details of general surgical procedures). A small hole was made in the bursa by tearing with sharp forceps, taking care to avoid blood vessels. Embryos to be implanted were transferred to a fresh drop of KSOM-AA without an overlay of oil, and then taken up into an oviduct transfer pipette attached to a mouth pipette assembly. The end of the oviduct transfer pipette was then inserted into the infundibulum that was being held with blunt forceps, and the embryos were expelled gently. Oviduct transfer pipettes were made the same way as uterine transfer pipettes (section 2.4.3), but pipettes with smaller bores were chosen.

2.4.7 LacZ staining of mouse embryos and adult tissues

Embryos were dissected from pregnant females on a specific embryonic day, determined by the presence of a copulation plug being E0.5. The uterus was removed from the abdominal cavity and placed in ice cold PBS contained within

a petri dish. Embryos were dissected away from the uterus, Reichert's membrane, amnion and placenta and transferred to fresh ice cold PBS for an additional 5 mins. Embryos were then fixed in 4% PFA for 15mins at 4°C, rinsed in PBS then incubated overnight at 28°C in X-gal stain base reagent (1mg/ml X-gal in N,N-Dimethylformamide, 2mM ($MgCl_2$), 5mM ($K_3Fe(CN)_6$), 5mM ($K_4Fe(CN)_6$), 0.24mM ($C_{24}H_{39}NaO_4$), in PBS, pH7.4-7.5). Following overnight staining the embryos were rinsed in PBS then post-fixed in 4% PFA overnight at 4°C. Embryos were rinsed for a final time in PBS then stored in 70% ethanol at 4°C. Tissues dissected from adult mice were fixed and stained in the same way as mouse embryos. Photographs were taken using a digital camera (RT. Color SPOT, Diagnostic Instruments Inc.) with embryos/tissues visualized under a dissection microscope (Leica) contained in a petri dish filled with PBS.

2.4.8 LacZ staining of sectioned adult tissues

Tissues were dissected from adult mice then any external moisture removed by gently dragging the tissue across foil. The dried tissues were then transferred into a cassette containing a layer of optimal cutting temperature compound (O.C.T.) embedding medium (Raymond Lamb), orientated as required and covered with a further layer of O.C.T. The cassettes were placed on top of dry ice to allow the O.C.T. to solidify then stored at -80°C for at least 24 hours prior to sectioning. Sections of 12 μm in thickness were cut using a cryostat (Leica, CM1850) and placed onto polysine coated slides (VWR International), dried at room temperature for 15 minutes then stored at -80°C for at least 24 hours prior to staining. All sectioning was performed by Iryna Withington, University of Bath. To stain the sections, slides were warmed at room temperature for 30 mins, a hydrophobic barrier pen (Vector Laboratories) was used to draw around the edge of each slide then 1ml of X-gal stain base reagent (1mg/ml X-gal in N,N-Dimethylformamide, 2mM ($MgCl_2$), 5mM ($K_3Fe(CN)_6$), 5mM ($K_4Fe(CN)_6$), 0.24mM ($C_{24}H_{39}NaO_4$), in PBS, pH7.4-7.5) was added to each slide, placed in a dark humidified chamber at 37°C overnight. The following day the slides were rinsed briefly in PBS, fixed in 4% PFA for 30 minutes, washed in an ascending ethanol series of 70%-100% for 1 min each to dehydrate the sections then washed twice in histoclear clearing solution (National Diagnostics) for 2 mins,

before finally being mounted using DPX mount (Sigma) overlaid with a coverslip.

2.4.9 Urine collection from mice

Urine was collected from mice either by scruffing and holding a 1.5ml microcentrifuge tube under the urethra or by individually housing the mice in cages without sawdust and using a pasteur pipette to collect urine from the bottom of the cage. The urine samples were immediately frozen until needed.

2.5 Specialist Materials Appendix

Anti-DIG AP fab fragments (Roche # 11093274910)

Borosilicate glass capillaries (Harvard apparatus #GC100T-10)

Colcemid (GibcoBRL #15210-040)

DIG RNA Labelling Kit (Roche #11175025910)

DNA-free kit (Ambion #1906)

DPX (Sigma #44581)

Dulbecco's modified eagle medium (Gibco BRL #61965-026)

ESGRO (10^7) leukaemia inhibitory factor (Chemicon International #ESG1107)

Foetal calf serum (Perbio #CH30160.03 batch #CLD0136)

Histoclear clearing solution (National Diognostics #HS-200)

Hydrophobic barrier pen (Vector Laboratories #H-4000)

MEM amino acids (Gibco BRL #11130-036)

MEM non-essential amino acids (Gibco BRL #1140-035)

Nylon transfer membrane (Osmonics #N00HYA0010)

O.C.T. (Raymond Lamb #LAMB/OCT)

Polysine coated slides (VWR International #031-0107)

Precast 10% Tris-HCl resolving gel (BIO-RAD #345-0010)

QIAshredder spin column (Qiagen # 79654)

RNeasy kit (Qiagen #74104)

XL1-Blue supercompetent cells (Stratagene #200249)

XL10-Gold® ultracompetent cells (Stratagene #200314)

Characterisation of the murine Basp1 gene

3.1 Summary

One of the aims of this thesis was to generate a mouse knockout model of the *Basp1* gene in order to study its role in development and its association with *Wt1* expression, particularly within the kidney. At the start of this project the sequence data for the mouse *Basp1* gene was not fully annotated within the sequence databases (e.g. Ensembl) and only limited information as to the murine expression pattern was known. It was therefore important to first find out about the sequence and structure of this gene and its expression pattern within the mouse.

3.2 Introduction

The *Basp1* gene is located at 5p15 within the human genome and on chromosome 15 within the mouse genome. The first aim of this chapter was to compare the human and mouse BASP1 protein and gene sequences. Whereas the human locus was well annotated, the mouse locus exhibited gaps where genomic sequence information was still unknown. Comparisons of the protein and gene sequences were carried out by local alignment algorithms then alignments were refined manually at the regions of similarity and discordancy. Murine genomic and cDNA sequences were also aligned to look for regions of overlap. Any cDNA sequences that did not overlap with the genomic sequence were investigated further by carrying out RT-PCR, to identify if they were truly part of a *Basp1* transcript or a sequencing error within the databases.

The second aim of this chapter was to study the expression profile of the *Basp1* gene during murine development, to add to and complement the limited expression data already known (see section 1.5). Expression data was compiled by RT-PCR, carried out on tissues of different developmental stages and by *in situ* hybridization of sectioned embryos at selected stages.

3.3 Methods

Primer sequences used for PCR or RT-PCR within this chapter can be found in Table 3.1. Plasmids used within this chapter can be found in Table 3.2.

3.3.1 Generation of a DNA probe for Northern hybridization

A *Basp1* DNA probe was made by digesting a plasmid containing a section of the murine *Basp1* gene (Figure 3.5A) with *BamHI* and *ApaI* followed by gel isolation of the ~425bp band containing the *Basp1* fragment.

3.3.2 Generation of riboprobes for *in situ* hybridization

Probes were made by digesting a plasmid containing a section of the murine *Basp1* gene (Figure 3.5A) with *BamHI* to generate an antisense (+ve) probe transcribed from T3 RNA polymerase, and with *NaeI* to generate a sense (-ve) probe transcribed from T7 RNA polymerase.

3.4 Results

3.4.1 Alignment of BASP1 protein sequences from different species

Protein alignments were carried out between the human protein sequence and the mouse, chicken, rat and cow protein sequences using a FASTA alignment with a SSEARCH function (Pearson, 1991; Smith and Waterman, 1981). Comparison of the human protein with the cow (76.6% identity within a 231 amino acid overlap), rat (70.3% identity within a 229 amino acid overlap) and mouse (71.7% identity within a 233 amino acid overlap) proteins were similar, while the chicken protein (63.1% identity within a 244 amino acid overlap) was more divergent (Figure 3.1).

3.4.2 Genomic and cDNA alignment

The human and rat *Basp1* transcripts contained two annotated exons, whereas the mouse genome contained only one. Aligning these regions using a BLAST alignment (Tatusova and Madden, 1999) showed that the single mouse exon overlapped with the second exon of the human genome, as did the two rat exons (Figure 3.2). Comparing mouse genomic sequences (e.g. AB046714) and cDNA

Table 3.1

primer pairs	sequences	Expected product size	annealing temp
3F 3R	$^{5'}$GCTCCAACTCGCTCCACGCTC$^{3'}$ $^{5'}$GGGTGGCTGGTGTCGTCCG$^{3'}$	92bp	60°C
2F 2R	$^{5'}$GGGCTACAATGTGAACGA$^{3'}$ $^{5'}$CTCGCTCTTCTCGGGCTCGG$^{3'}$	255bp	60°C
3F 2R	$^{5'}$GCTCCAACTCGCTCCACGCTC$^{3'}$ $^{5'}$CTCGCTCTTCTCGGGCTCGG$^{3'}$	429bp	60°C
2FN 4R	$^{5'}$GCAAAGCGTGCCGTCAAAGAG$^{3'}$ $^{5'}$CCCTTGGGCAATACATATCCTCAC$^{3'}$	172bp	60°C
4F 5R	$^{5'}$TTGGAAGTGAGGATATGTATTGCC$^{3'}$ $^{5'}$TGGGCTCCGTCTGAAAGTTGG$^{3'}$	443bp	60°C
5F 6R	$^{5'}$TGCCAACTTTCAGACGGAGCC$^{3'}$ $^{5'}$TGACCAATAAACAGAACTACAGG$^{3'}$	186bp	60°C
5F 7R	$^{5'}$TGCCAACTTTCAGACGGAGCC$^{3'}$ $^{5'}$GTTCATCTTAGGGGATAGTGGTC$^{3'}$	583bp	60°C
6F 7R	$^{5'}$AACATTCTCTCATCATACTTAGCC$^{3'}$ $^{5'}$GTTCATCTTAGGGGATAGTGGTC$^{3'}$	490bp	60°C
7F 8R	$^{5'}$ACCACTATCCCCTAAGATGAACA$^{3'}$ $^{5'}$AAGCCAAAGCAAGTAGGTCTG$^{3'}$	331bp	60°C
8F 9R	$^{5'}$CAGACCTCACTTGCTTTGGCTT$^{3'}$ $^{5'}$TGGGCATTTCTCTCCGATTCAG$^{3'}$	493bp	60°C
9F 10R	$^{5'}$CTGAATCGGAGAGAAATGCCCA$^{3'}$ $^{5'}$ATAGTTGGCTGGTGGCTAATGC$^{3'}$	385bp	60°C
10F 11R	$^{5'}$AAAGCATTAGCCACCAGCCAAC$^{3'}$ $^{5'}$TTTAGTTGTTCCATTGCCTGCCT$^{3'}$	561bp	60°C
11F 12R	$^{5'}$CAAGGCAGGCAATGGAACAACT$^{3'}$ $^{5'}$CTCACCAATCCTGAAACAATGCC$^{3'}$	502bp	60°C

12Fii 13Rii	$^{5'}$GCATCAAGTCTCTGCCAAAGC$^{3'}$ $^{5'}$GGGGACCTGCCAGTTCAAATG$^{3'}$	600bp	60°C
13F 14R	$^{5'}$CATACAGAGCAGGCTGGTTTCA$^{3'}$ $^{5'}$GTGTTCACGGATACAATGGCAC$^{3'}$	570bp	60°C
14Fii 15Rii	$^{5'}$TAACACTGAGAAACAACACGGC$^{3'}$ $^{5'}$TAATCTGTCCCTCGCCTCCTG$^{3'}$	496bp	60°C
14F 15R	$^{5'}$GAGTGCCATTGTATCCGTGAAC$^{3'}$ $^{5'}$TTTATGTGAACTACCCGCCCCT$^{3'}$	589bp	60°C
13Fiilong	$^{5'}$CCAAATGCTGGGACTAAAGGCACGTA$^{3'}$	RACE primer	

Table 3.1. Primer pairs used to perform PCR or RT-PCR. Shows primer names, sequences, the size product expected and the annealing temperature of the primer pairs.

Table 3.2

resource/plasmid name	**reference/source**	**description**
mGapdh probe	Nomura et al., 1989	280bp of the murine *Gapdh* gene cDNA cloned into pGEM3Zf, digested with *HindIII* and *PstI*.

Table 3.2. Details of the plasmid used to generate the mGapdh probe.

Figure 3.1. Alignment of the human (upper) and mouse (lower) Basp1 protein sequences. Two dots indicate an exact amino acid match; one dot indicates similar amino acids; no dot indicates dissimilar amino acids; a dash indicated where a gap had to be inserted in order to achieve the best alignment.

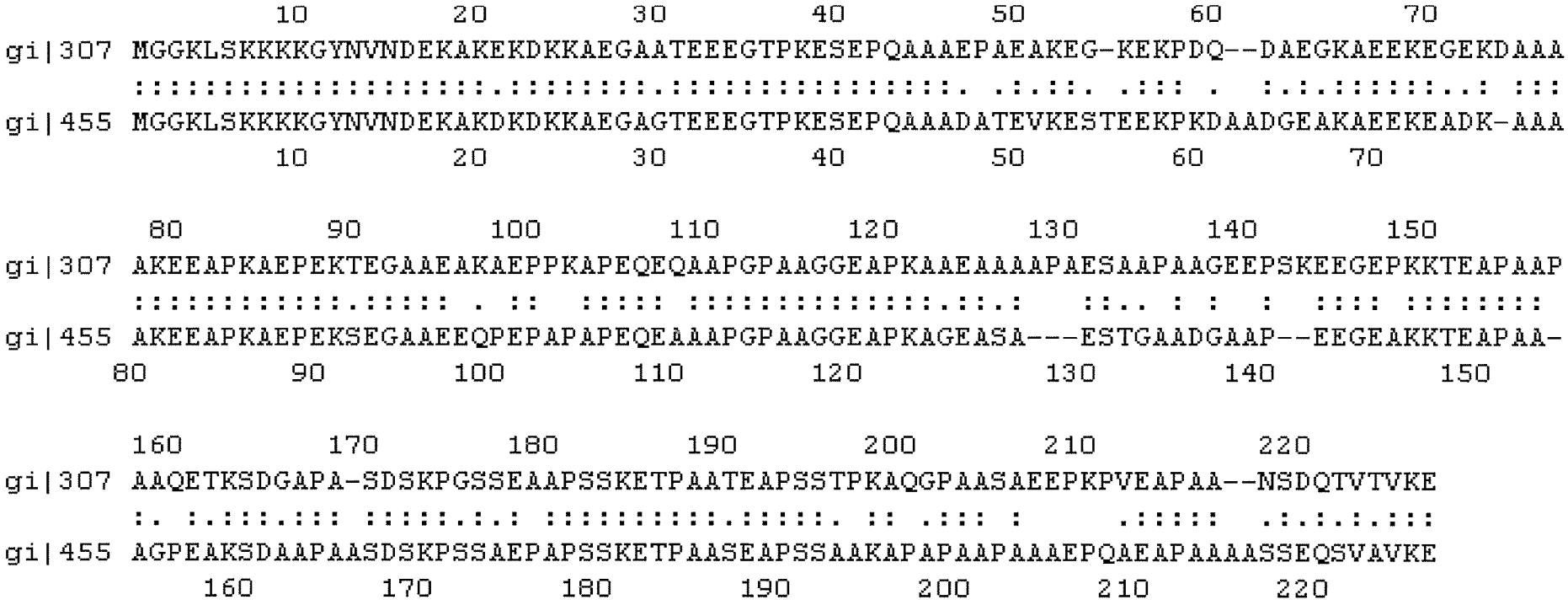
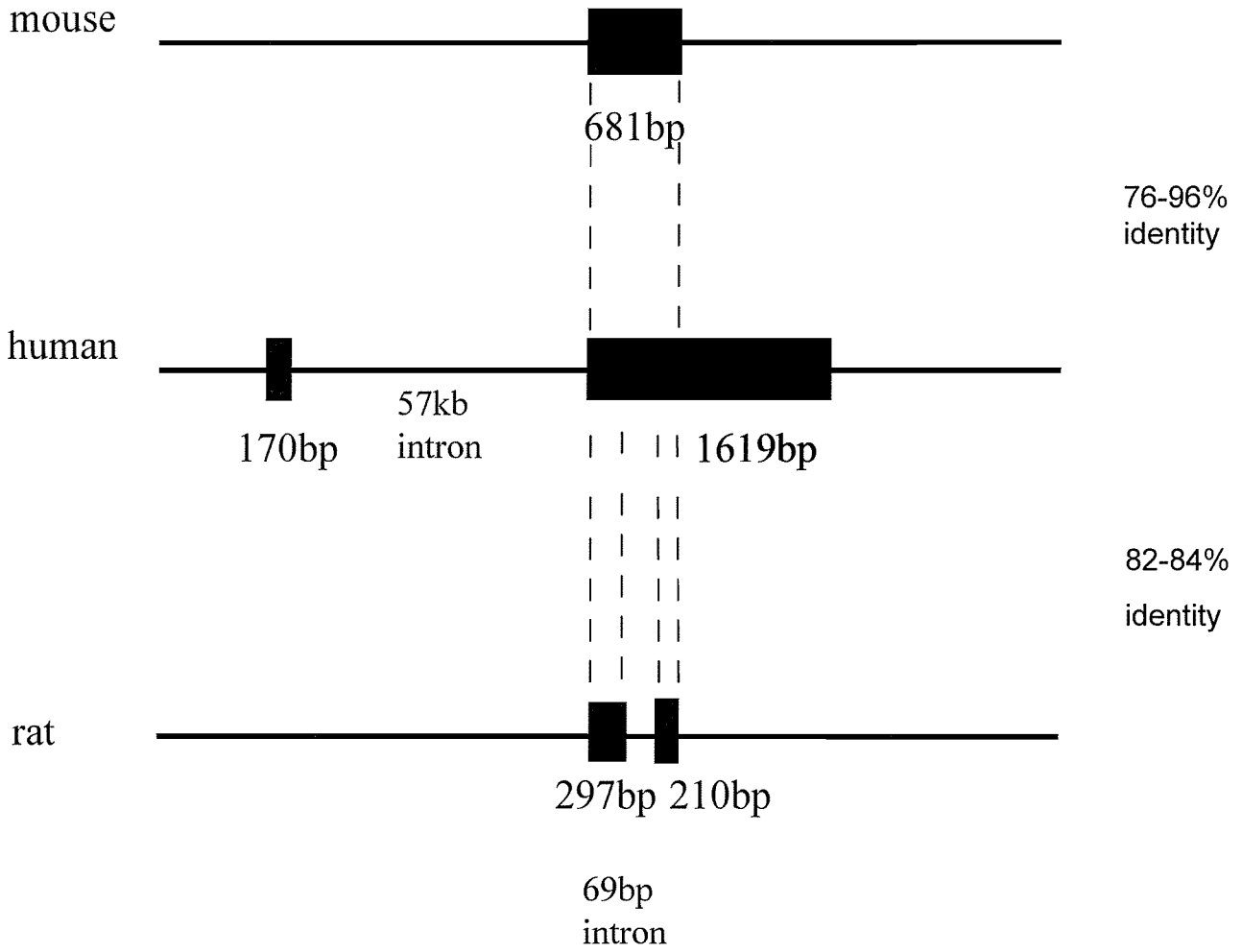

10 20 30 40 50 60 70
gi|307 MGGKLSKKKKGYNVNDEKAKEKDKKAEGAATEEEGTPKESEPQAAAEPAEAKEG-KEKPDQ--DAEGKAEEKEGEKDAAA
gi|455 MGGKLSKKKKGYNVNDEKAKDKDKKAEGAGTEEEGTPKESEPQAAADATEVKESTEEKPKDAADGEAKAEEKEADK-AAA
10 20 30 40 50 60 70
80 90 100 110 120 130 140 150
gi|307 AKEEAPKAEPEKTEGAAEAKAEPPKAPEQEQAAPGPAAGGEAPKAAEAAAAPAESAAPAAGEEPSKEEGEPKKTEAPAAP
gi|455 AKEEAPKAEPEKSEGAAEEQPEPAPAPEQEAAAPGPAAGGEAPKAGEASA---ESTGAADGAAP--EEGEAKKTEAPAA-
80 90 100 110 120 130 140 150
160 170 180 190 200 210 220
gi|307 AAQETKSDGAPA-SDSKPGSSEAAPSSKETPAATEAPSSTPKAQGPAASAEEPKPVEAPAA--NSDQTVTVKE
gi|455 AGPEAKSDAAPAASDSKPSSAEPAPSSKETPAASEAPSSAAKAPAPAAPAAAEPQAEAPAAAASSEQSVAVKE
160 170 180 190 200 210 220

Figure 3.2. Schematic representation of the human, mouse and rat *Basp1* genomic regions showing the position of the annotated exons as they appeared at the outset of this project. The regions of shared sequence identity and the overall percentage identity of the sequences between species are indicated.


mouse
681bp
76-96%
identity
human
57kb
intron
170bp
1619bp
82-84%
identity
rat
297bp 210bp
69bp
intron

sequences (e.g. AK011545) revealed a potential upstream exon due to the 5' cDNA region not aligning with the genomic region immediately upstream of the annotated exon. Additionally, the 3' end of the cDNA sequence did not end at the same position as the annotated exon within Ensembl, although it did align with the genomic sequence, suggesting that this exon was larger than that shown on the sequence annotation. The sequence identity between the first annotated exon of the human *BASP1* gene and the potential upstream exon of the mouse *Basp1* gene was 83%.

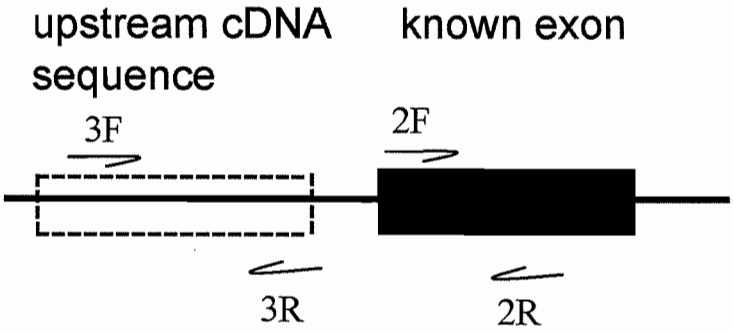
3.4.3 Finding an additional exon in the mouse *Basp1* sequence

To investigate if there really was an upstream mouse exon, RT-PCR was carried out using ES cell RNA with primers designed to the known exon [2F/2R], the 5' upstream cDNA sequence [3F/3R] and between the two regions [3F/2R] (Figure 3.3). The size product expected with primers 3F/2R, if the 5' upstream region was an exon immediately upstream of the annotated exon, was 429bp. The results of these RT-PCRs confirmed the presence of a second exon within the mouse genome. This information was essential to the design of the targeted *Basp1* knockout mouse (see Chapter 4).

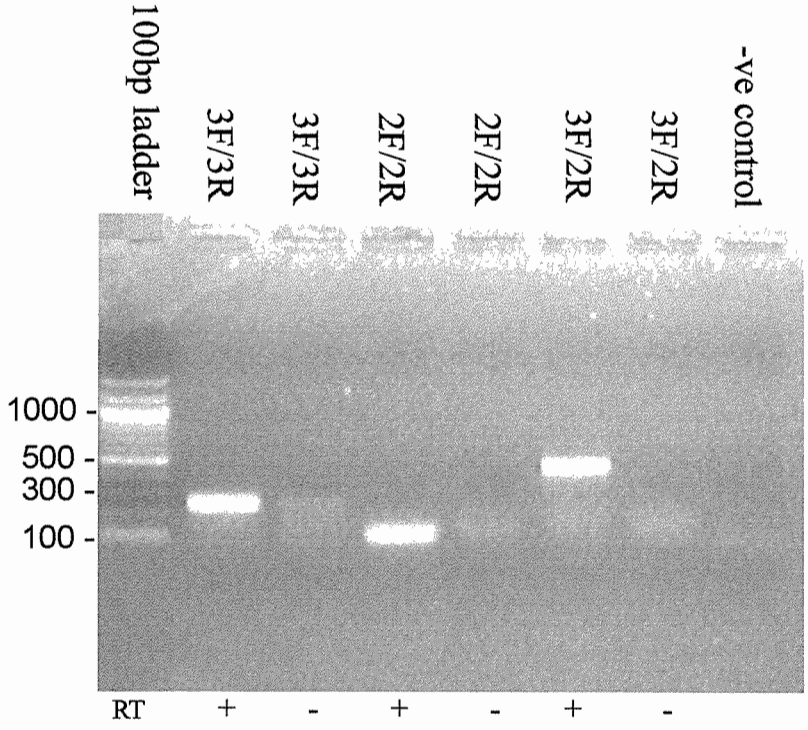
Due to the mouse *Basp1* gene not being fully annotated within the sequence databases, searching for the exon 1 genomic sequence by BLAST did not find a match to the region. It was therefore not possible to ascertain the size of the intronic region between exons 1 and 2. In order to find the size of this intronic region, PCR was carried out on genomic DNA extracted from a mouse tail tip with primers designed within exon 1 and exon 2 using an approach that should yield long template PCR products of up to 6kb. If the size of this intron within the murine locus was as long as in the human locus (57kb), it would not have been possible to amplify such a large region using this method. No PCR product was amplified across the intron (data not shown) which was not surprising given that subsequently the sequence databases have been updated to include two mouse exons with an intron size of 48.6kb. As the size of the intron was not essential to the design of the targeting construct to generate a knockout mouse,

Figure 3.3. RT-PCR to verify the existence of a second *Basp1* exon, 5' of the annotated exon within the sequence databases. A, schematic of the *Basp1* genomic region showing the position of the annotated exon, the presumed upstream exon and the position of primers used in RT-PCR experiments. B, agarose gel showing RT-PCR products obtained with different primer pairs using cDNA made from ES cell RNA. Lanes represent RT-PCR reactions carried out with first strand synthesis reactions made with (RT +) or without (RT -) reverse transcriptase.

A


upstream cDNA sequence
known exon
3F
2F
3R
2R

B


100bp ladder
3F/3R
3F/3R
2F/2R
2F/2R
3F/2R
3F/2R
-ve control
1000 -
500 -
300 -
100 -
RT + - + - + -

no further attempts to find out this information were attempted prior to the sequence databases being updated.

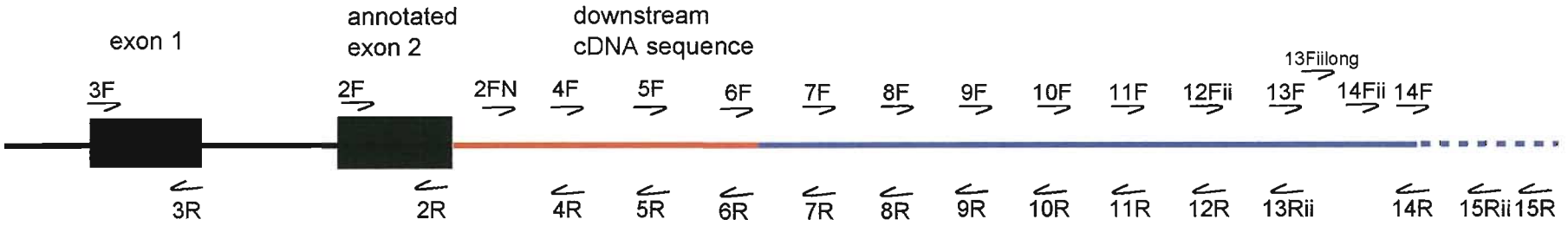
3.4.4 Searching for the 3' end of the second exon of the mouse *Basp1* gene

In order to find the size of exon 2 of the murine *Basp1* gene, RT-PCR was carried out using ES cell RNA with primers designed downstream of the 3' end of the annotated exon within Ensembl (Figure 3.4). An RT-PCR 'walk' was first carried out with successive primer pairs being designed downstream of the characterized exon within the 3' cDNA region (red line in Figure 3.4A). RT-PCR products were generated for all primer pairs designed within this 3' cDNA region. Next, successive primer pairs were designed downstream of the 3' cDNA region within the genomic sequence (blue line in Figure 3.4A). Within the downstream genomic region, RT-PCR products were generated with all primer pairs (apart from 8F/9R) up until primers 13F/14R when no products were obtained after this point. It was unlikely that the RT-PCRs did not work due the failure of primers, as two sets of primer pairs (14F/15R and 14Fii/15Rii) designed 3' of 13F/14R were used. The size of exon 2 was therefore estimated to be roughly 4.8kb. However, if a splice was present spanning the region of 3'-most primers, this would provide another reason for the RT-PCR not working apart from being the 3' end of the transcript.

In order to more precisely locate the 3' end of exon 2, 3' RACE PCR was carried out using nested gene specific primers 13F and 13Fiilong. Five PCR products could be seen when run out on an agarose gel (~ 200bp, 450bp, 500bp, 900bp, 1kb) (data not shown). Due to information obtained from the RT-PCR 'walk', it was known that the RACE product should be between 500bp and 1.1kb, therefore, the largest three bands were isolated from the gel and cloned into pGEM T-easy for sequencing. To select for colonies that could potentially contain the correct fragment, restriction digests were carried out on plasmid DNA extracted from several colonies choosing restriction sites that were known to be adjacent to the gene specific primers used. None of the colonies containing the 500bp inset and the 1kb insert gave restriction digest patterns of the expected size, whereas the colonies containing the 900bp insert did. Colonies containing

Figure 3.4. RT-PCR to find the 3' end of *Basp1* exon 2. A, schematic of the genomic locus showing the position of primers. Black boxes indicate the known exonic sequences; red line indicates the downstream cDNA region; blue line indicates the downstream genomic region. B, table showing which primer pairs gave RT-PCR products. C, agarose gel showing RT-PCR products obtained with different primer pairs using cDNA made from ES cell RNA. Not all primer pairs are represented. Lanes represent RT-PCR reactions carried out with first strand synthesis products made with (RT +) or without (RT -) reverse transcriptase, or a water control (0).

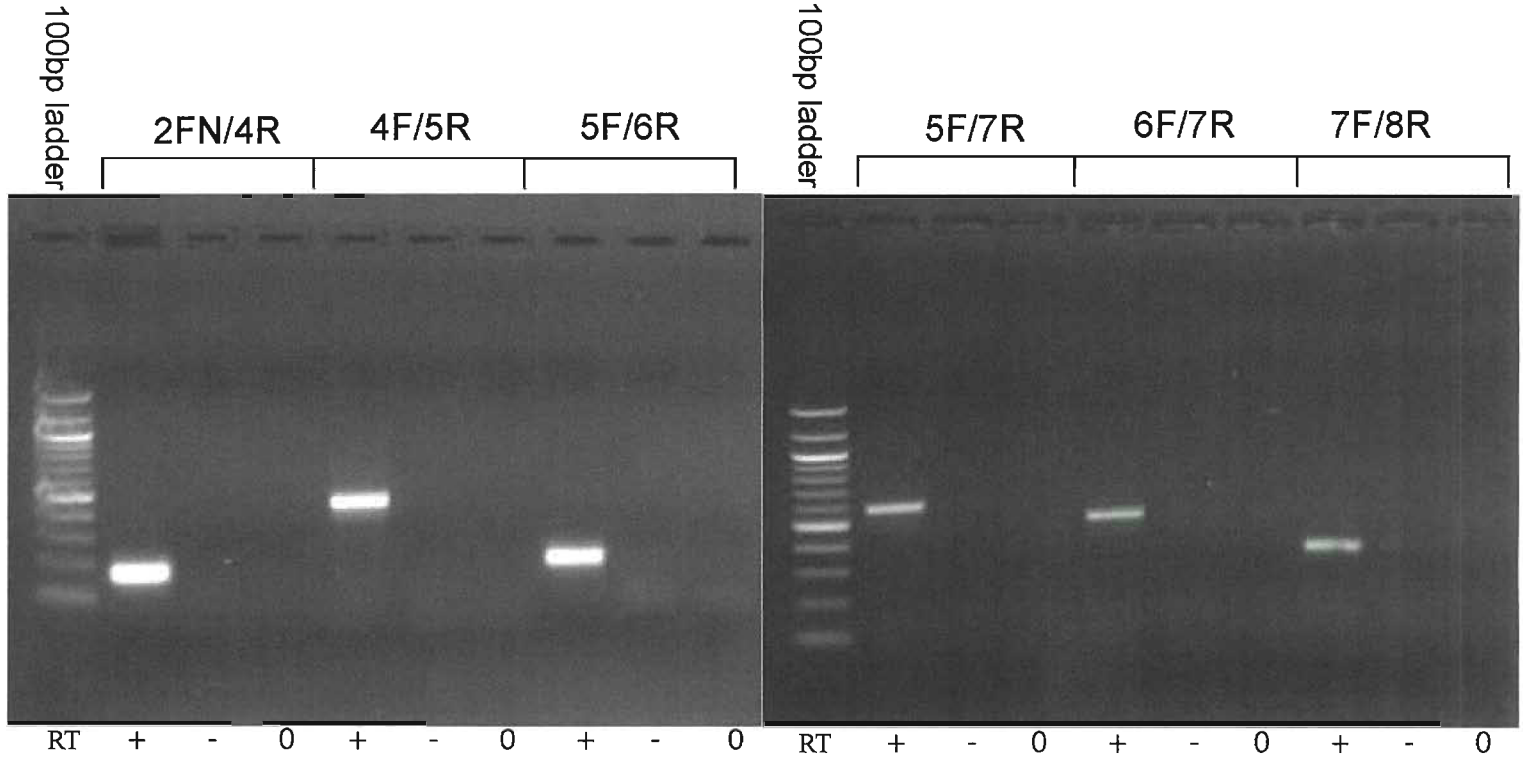
A



B

primer pairs	RT-PCR product
2FN/4R	✓
4F/5R	✓
5F/6R	✓
5F/7R	✓
6F/7R	✓
7F/8R	✓
8F/9R	×
9F/10R	✓
10F/11R	✓
11F/12R	✓
12Fii/13Rii	✓
13F/14R	✓
14Fii/15Rii	×
14F/15R	×

C



the 900bp insert were sequenced and compared to the *Basp1* locus by BLAST searching; no significant matches were found. As it was not imperative that the exact 3' end of the mouse *Basp1* gene was found, Northern hybridizations were carried out next to find the entire length of the RNA transcript.

3.4.5 Determination of the *Basp1* RNA transcript length by Northern hybridization

Due to the annotation within the sequence databases of a novel transcript overlapping with the murine *Basp1* gene and transcribed in the opposite direction, Ensemble v38; entry **ENSMUST00000090276**, it was possible that this gene had an antisense transcript. In an attempt to verify the existence of an antisense transcript and to find the size of the *Basp1* sense RNA transcript, Northern hybridization was carried out using strand-specific riboprobes (see section 2.1.17), however, even following stringent washing, due to non-specific binding the only hybridization products that could be seen were the 28S and 18S ribosomal bands. The next approach taken was to carry out a Northern hybridization with DNA probes, the advantage being that the protocol was more robust, the disadvantage being that if more than one hybridization band was seen it would not be possible to distinguish between a sense and antisense RNA transcript. A *Basp1* DNA probe (see section 3.3.1) was used to probe a blot of total RNA from adult brain. Only one hybridization product of between 2-3kb compared to the RNA ladder standard was seen following the Northern protocol (Figure 3.5B). To verify this size, the blot was stripped and re-probed with a murine *Gapdh* (mGAP) probe (Nomura et al., 1989) (see Table 3.2) that gives a known hybridization size product of 1.24kb (Figure 3.5B). Comparing the *Gapdh* hybridization product to the *Basp1* hybridization product, suggested the *Basp1* transcript size was likely to be closer to 2kb than 3kb.

3.4.6 RT-PCR analysis of *Basp1* expression at different developmental stages

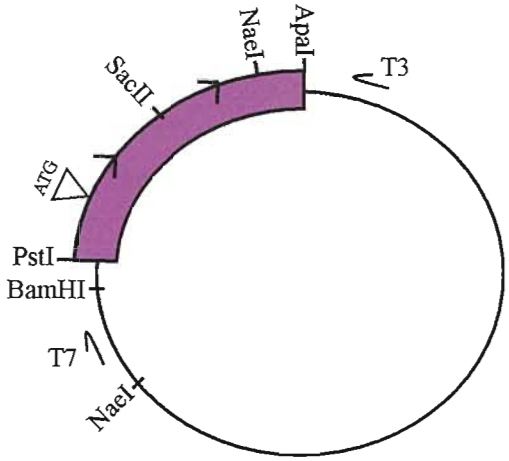
In order to find out at what mouse developmental stages the *Basp1* gene was expressed, total RNA was extracted from a series of staged tissues (E13.3 head, E13.3 trunk E13.3 placenta, E15.5 kidney, E15.5 placenta, E17.5 kidney, E18.5

Figure 3.5. *Basp1* RNA transcript size determined by Northern hybridization. A, schematic of the plasmid used to generate the DNA probes for Northern hybridization. B, hybridization products generated with the *Basp1* probe (left hand side).and the *Gapdh* probe (right hand side). The sizes were estimated using the Fermentas High Range RNA ladder.

Figure 3.6. Agarose gel showing RT-PCR products to determine at which developmental stages *Basp1* is expressed. Lanes represent RT-PCR reactions carried out with first strand synthesis products made with (RT +) or without (RT -) reverse transcriptase.

Figure 3.5

A



B

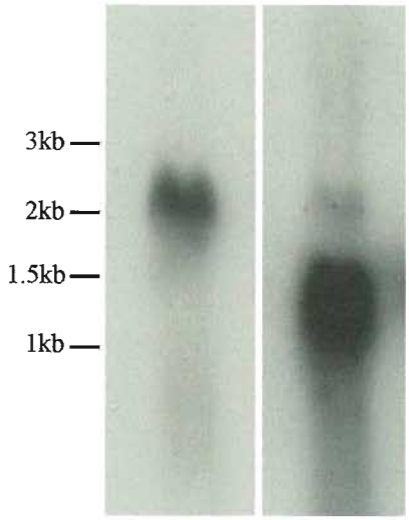
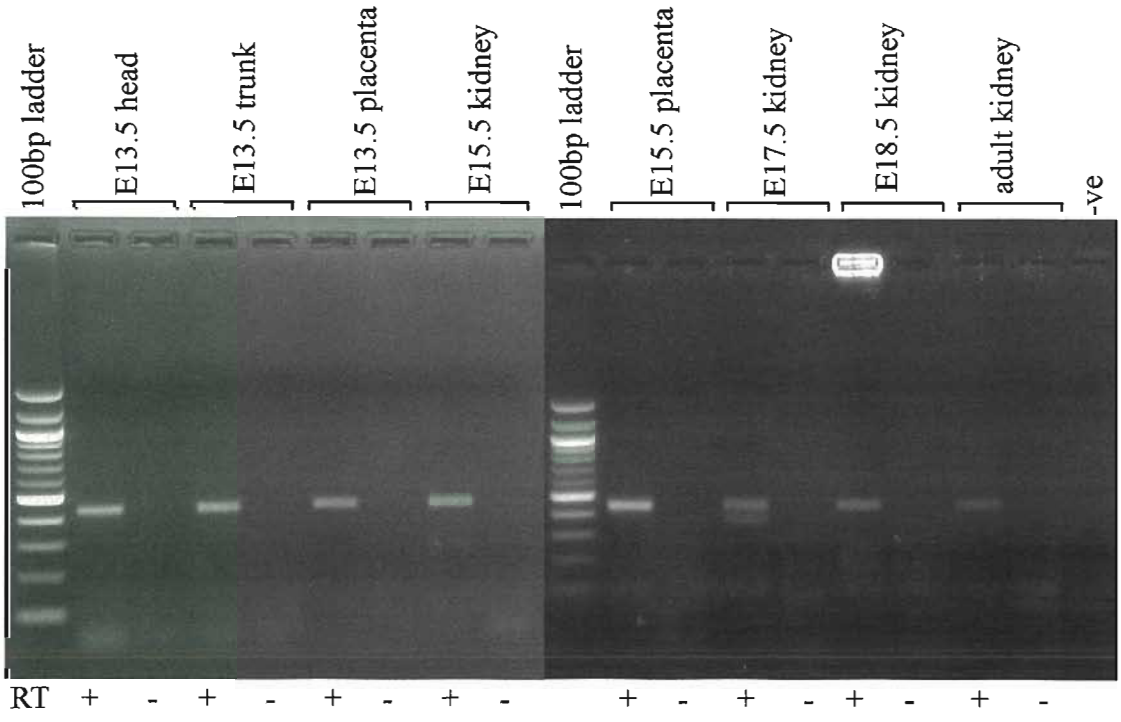


Figure 3.6



kidney and adult kidney) so that RT-PCR analysis could be carried out. OligodT-primed cDNA was made and RT-PCR carried out with primers 3F/2R that span exons 1 and 2 (Figure 3.6). *Basp1* was expressed in each of these staged tissues. The smaller sized band obtained with the E17.5 kidney sample was most likely to be due to a PCR artefact as this product was not amplified when the RT-PCR was repeated for this staged tissue.

3.4.7 *In situ* hybridization of sectioned embryos

In order to find out more about the expression pattern of *Basp1* during murine development, *in situ* hybridization was carried out on sectioned embryos and adult tissues using a *Basp1 in situ* probe.

E15.5

High levels of *Basp1* RNA expression detected by the antisense probe could be seen in the trigeminal ganglion, thyroid gland, dorsal root ganglion, retina, intestines, nephrons of the kidney, testis, liver, lung, salivary glands, a structure that may be a tooth and several regions of the brain including the neopallial cortex and ventricular zone of the telencephalic vesicle of the brain, the roof of the dorsal midbrain and the choroid plexus arising from the fourth ventricle (Figure 3.7A and B). Lower levels could be seen in the dermis and spinal cord.

E13.5

Similar to E15.5, high levels of *Basp1* expression were seen in several regions of the brain, epiglottis, dorsal root ganglion, the intestines and the lung (Figure 3.7A and C). There was stronger expression in the spinal cord and dermis and also regions not represented in the E15.5 *in situ* sections, the cartilage of the jaw, tongue, olfactory epithelium and cartilage of the tail/limb. There was little if any expression in the liver, although the *in situ* at this developmental stage was carried out only once.

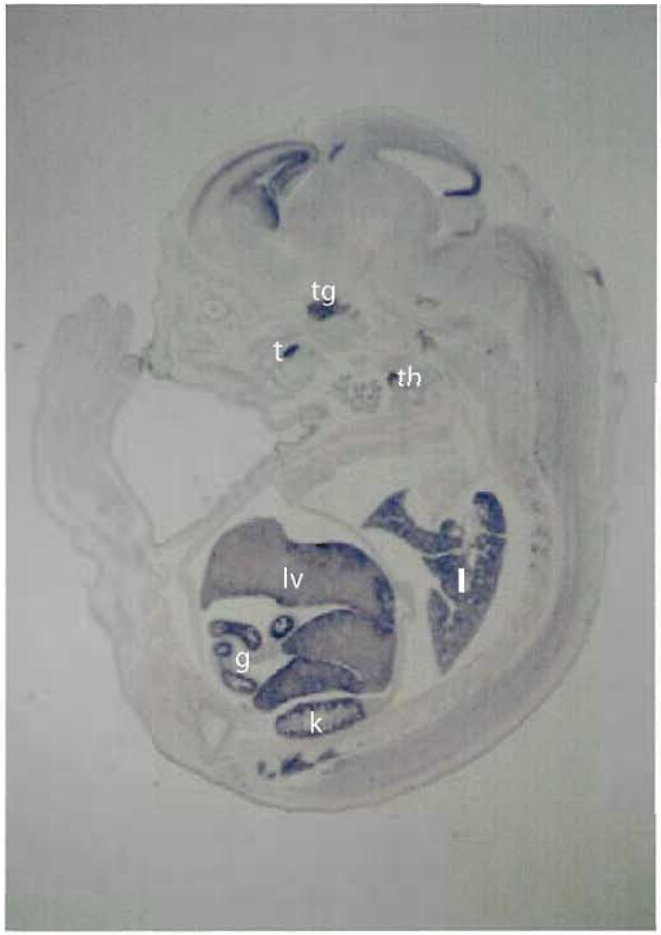
E10.5

At this developmental stage, few organs have formed. High levels of *Basp1* expression were seen throughout the embryo including the wall of the brain, forelimb bud, branchial arch and neural tube (Figure 3.7A and D). Interesting, unlike the other developmental stages observed, expression was high in regions of the developing heart, including the lining of the atrial chamber and the bulbus cordis, a transitory structure of the embryonic heart.

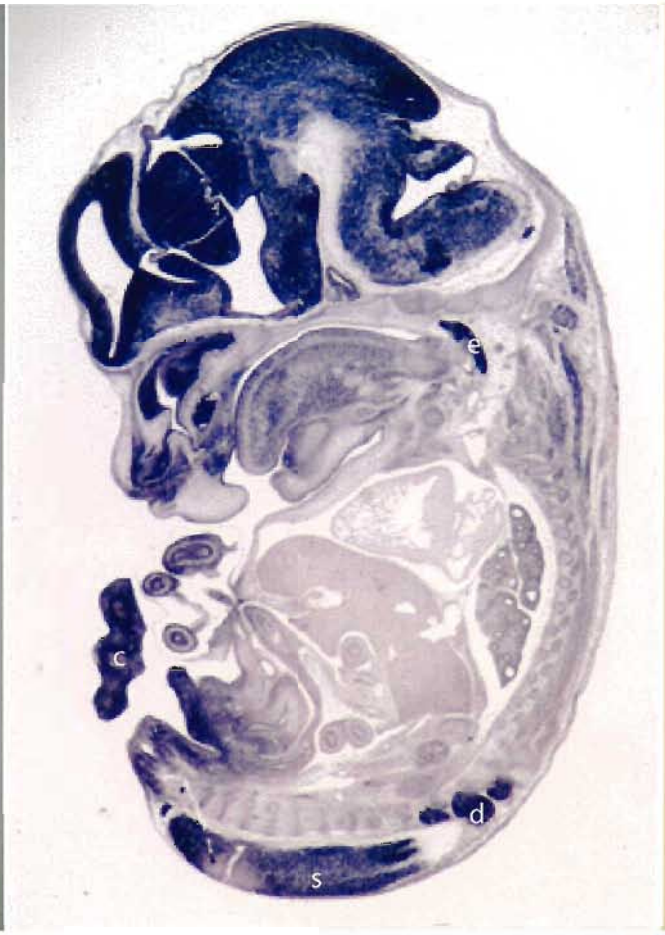
Figure 3.7. *In situ* hybridisation photographs of mouse embryos and mouse adult organs. A, whole embryo sections hybridized with a *Basp1* antisense probe. i, E15.5 (10x); ii, E13.5 (16x); iii, E10-E10.5 (12.5x). lv, liver; l, lung; g, intestine; k, kidney; tg, trigeminal ganglion; s, spinal cord; d, dorsal root ganglion; e, epiglottis; c, cartilage; b, bulbus cordis; a, atrial chamber; th, thyroid gland; t, tooth. B-E, *Basp1* antisense probe (left hand panels) and *Basp1* sense probe (right hand panels). B, magnification of E15.5 organs within whole embryo sections. Panels i, ii neopallial cortex and ventricular zone of the telencephalic vesicle of the brain; iii, iv, roof of dorsal midbrain; v, vi, choroids plexus arising from the roof of the lateral recess of the fourth ventricle of the brain; vii, viii, testis; ix, x, dorsal root ganglion; xi, xii, dermis; xiii, xiv, inner neural layer of the retina of the eye; xv, xvi, intestinal epithelium; xvii, xviii, developing nephrons of the kidney; xix, xx, liver; xxi, xxii, lung; xxiii, xxiv, salivary glands. Photographs shown at 200x magnification. C, magnification of E13.5 organs within whole embryo sections. Panels i, ii, jaw and tongue; iii, iv, nasal cavity; v, vi, lung; vii, viii, mesentery of the midgut. Photographs shown at 200x magnification. j, jaw; t, tongue. D, magnification of E10.5 organs within whole embryo sections. Panels i, ii, atrial chamber; iii, iv, bulbus cordis. Photographs shown at 200x magnification. E, adult organs. Panels i, ii, kidney; iii, iv, testis. Photographs shown at 200x magnification. Organs were identified with the aid of 'The atlas of mouse development' by Kaufman, 1994.

A

i

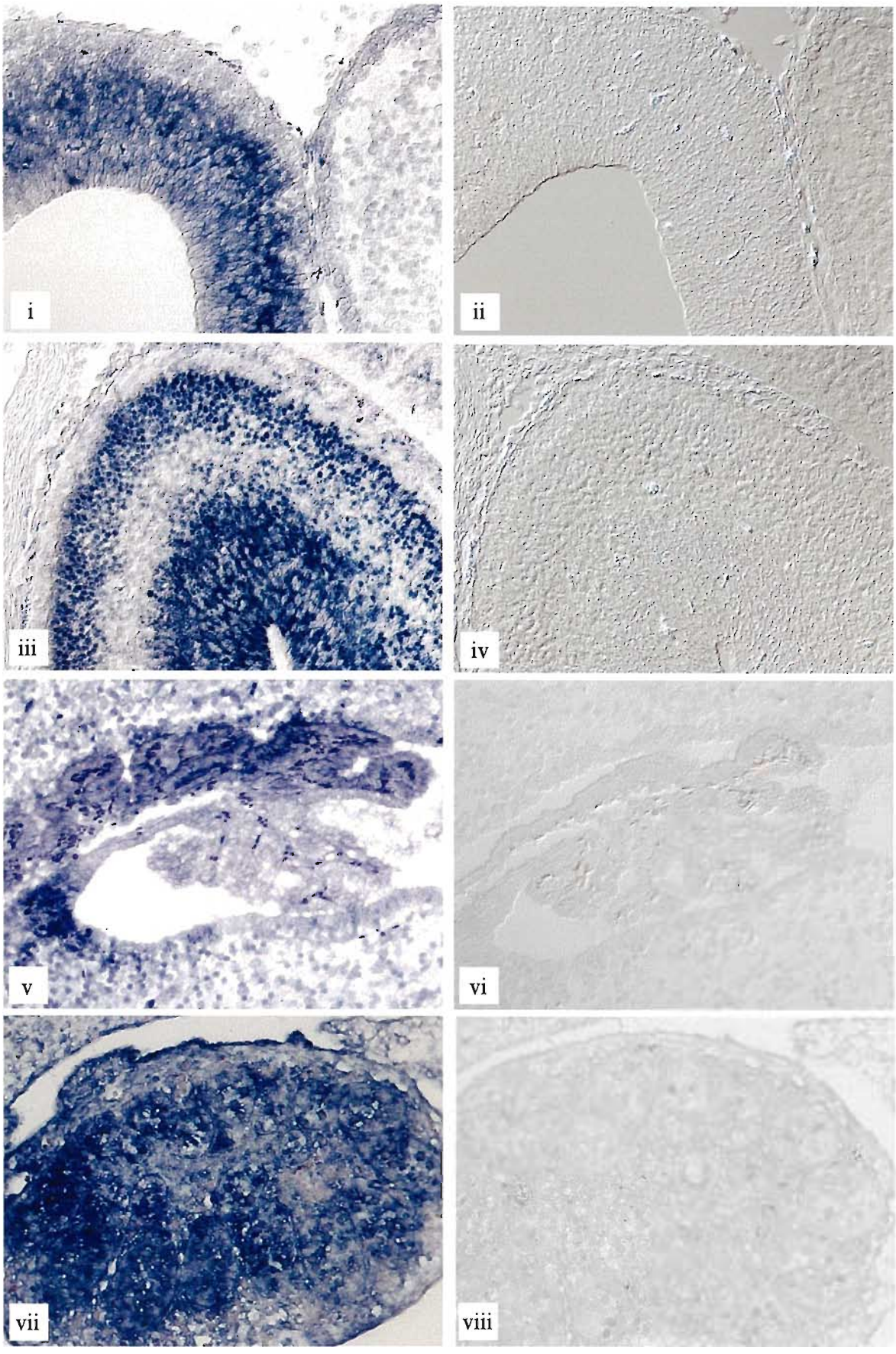


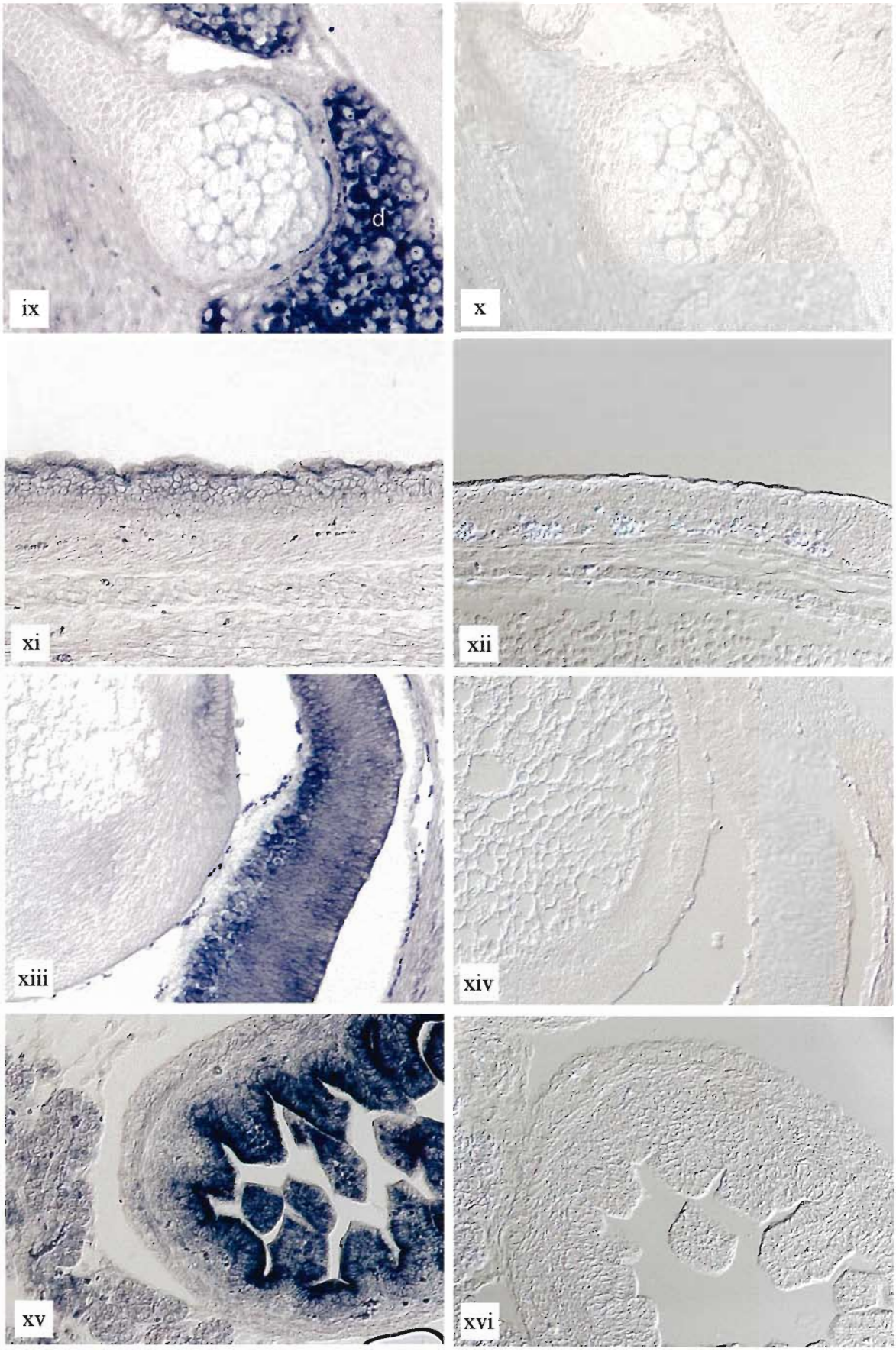
ii

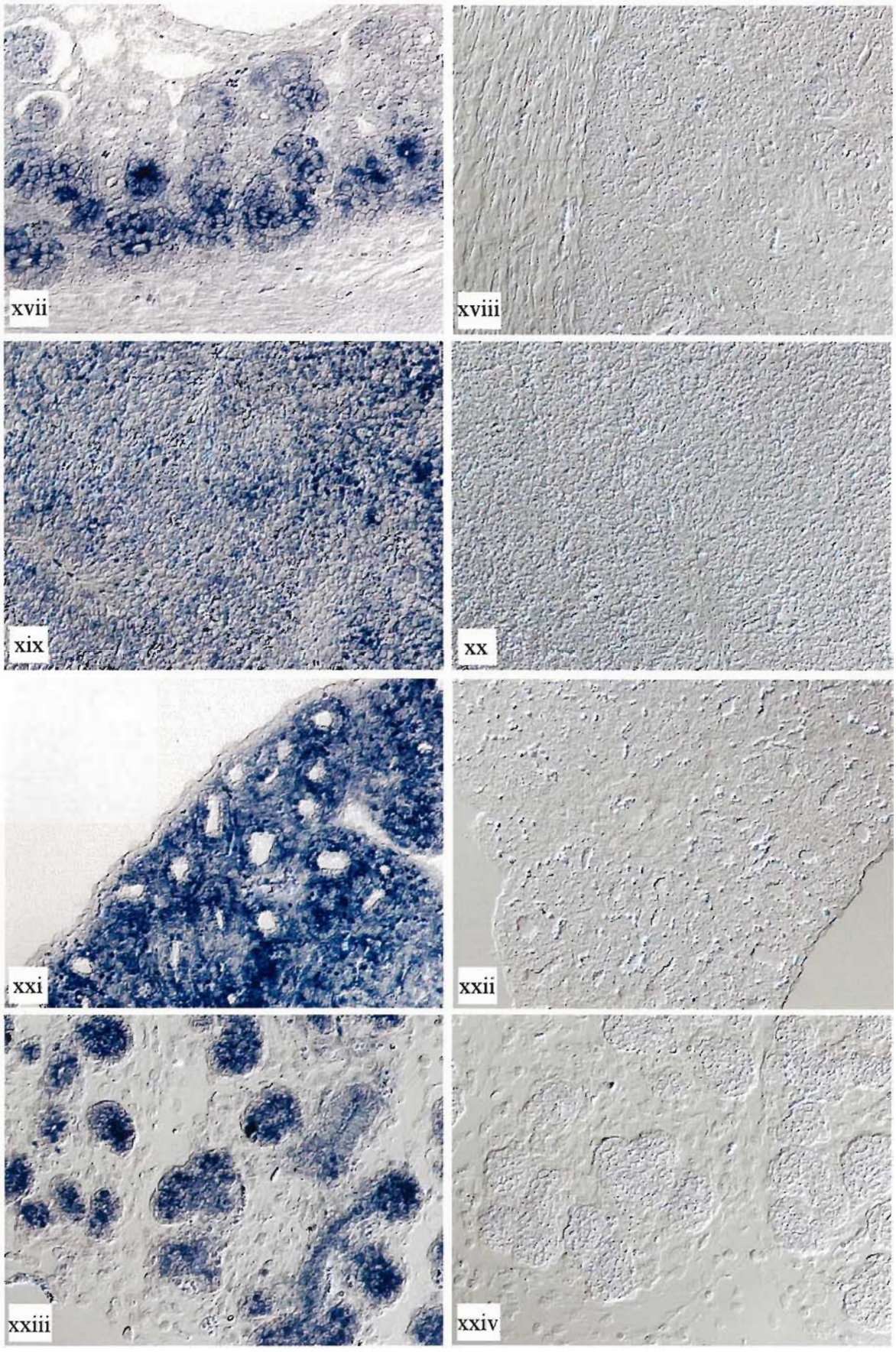


iii

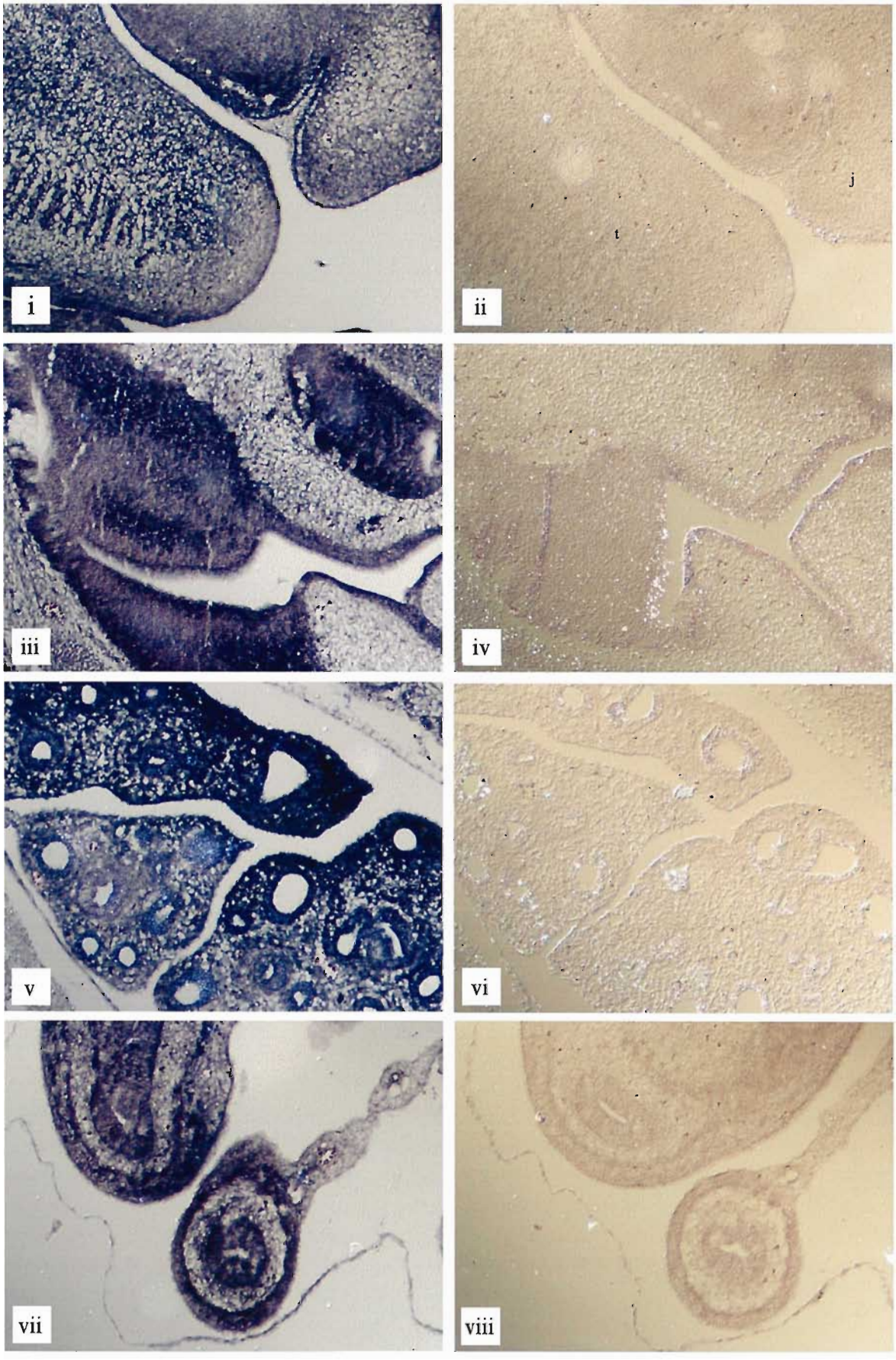
B



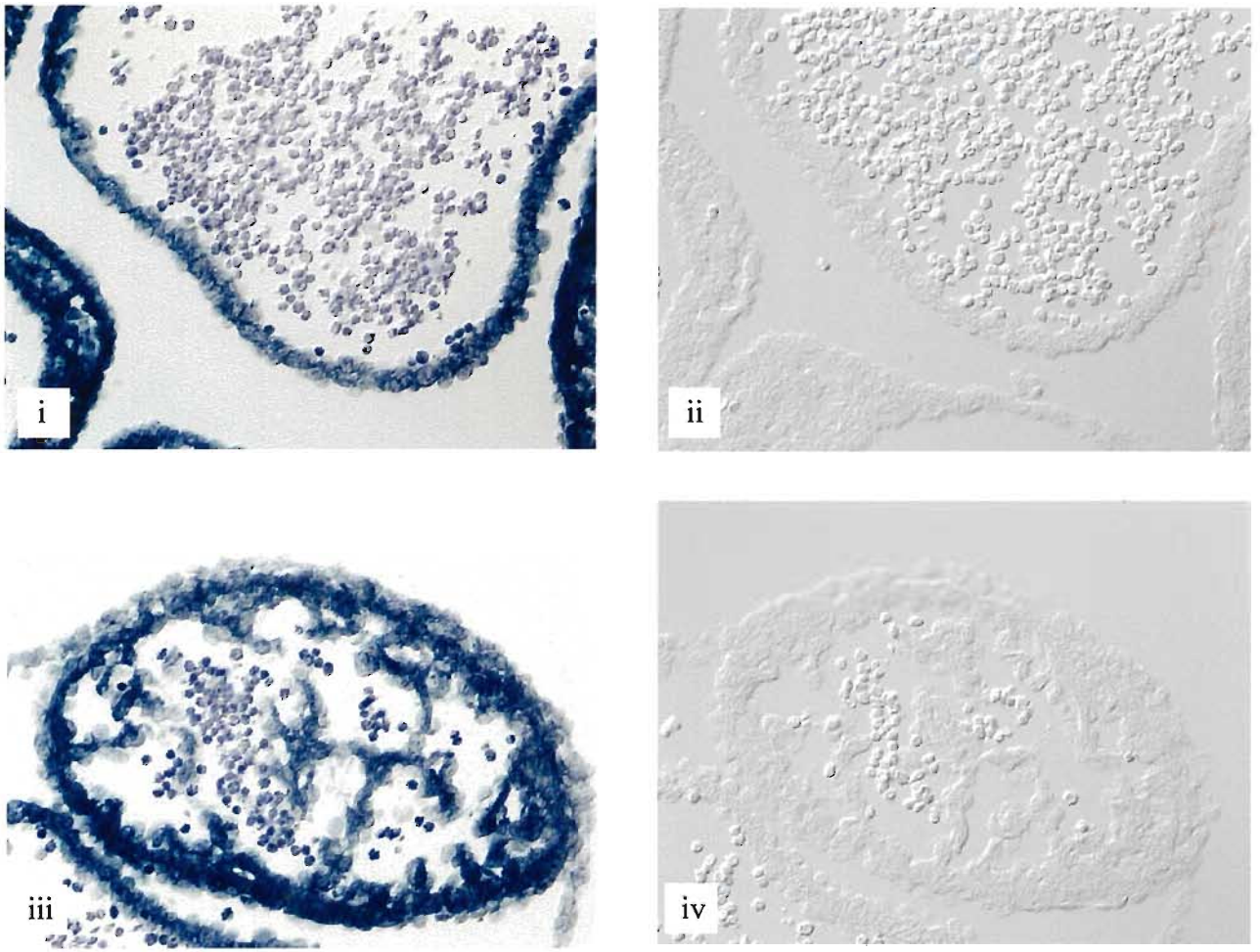

d
ix
x
xi
xii
xiii
xiv
xv
xvi


xvii
xviii
xix
xx
xxi
xxii
xxiii
xxiv

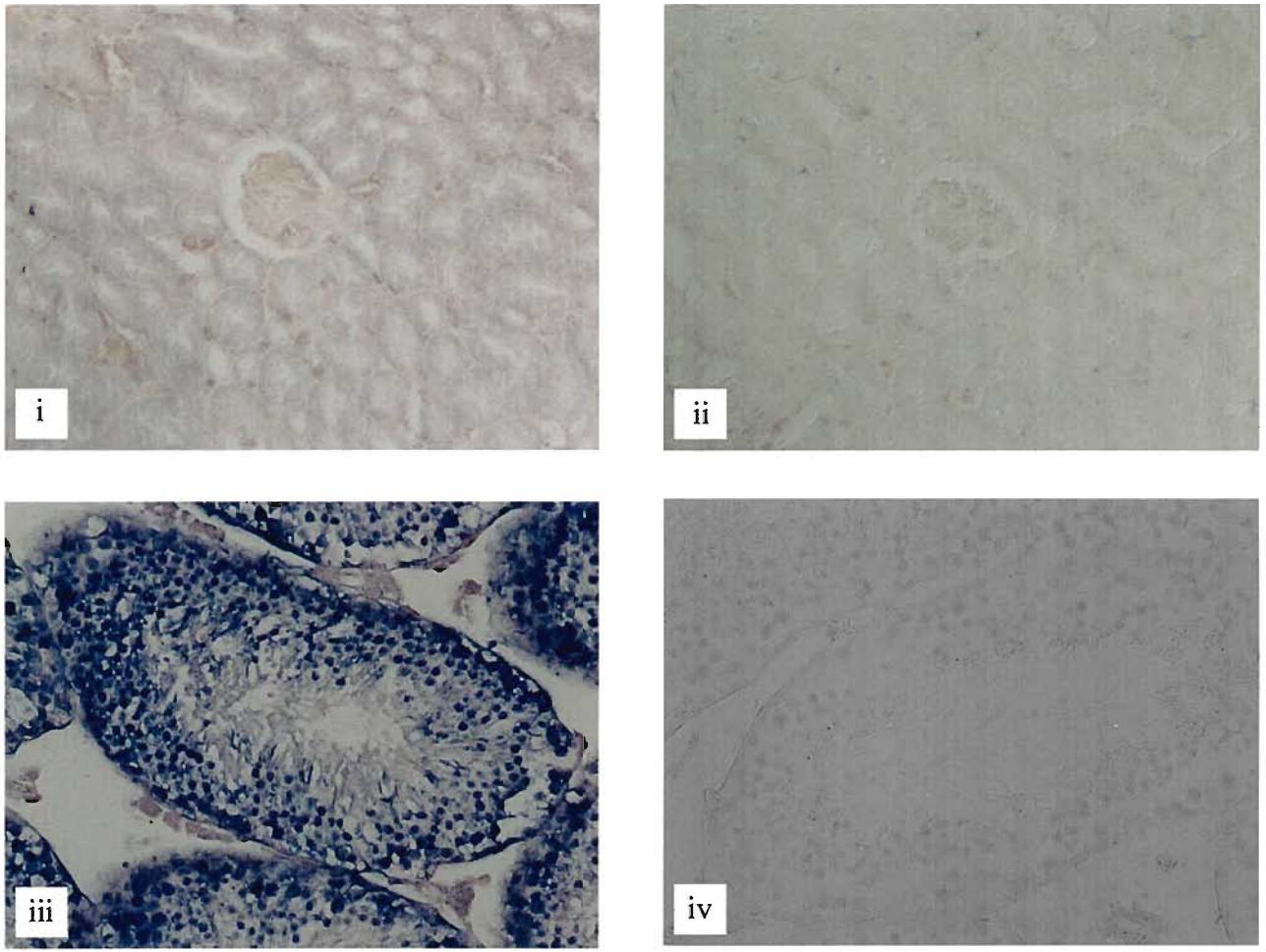
C



D



E



Adult kidney

No specific *Basp1* expression was seen within the adult kidney (Figure 3.7E). There were regions of intense staining which appeared in sections with both the sense and antisense probes, namely the epithelial lining of the kidney and tissue lining the blood vessels, but this was put down to non-specific binding of the probes.

Adult testis

High levels of *Basp1* expression were seen in the cells within the outer layer of the seminiferous tubules, presumably within the Sertoli cells, as the mature gamete cells within the central region of the seminiferous tubules have no *Basp1* expression suggesting that the immature gamete cells, which were difficult to distinguish from the Sertoli cells, do not express *Basp1* either (Figure 3.7E).

3.5 Discussion

In order to further characterise the role of *Basp1* during murine development, the expression pattern of this gene has been investigated. Local sequence alignments of *Basp1* between species along with molecular evidence has also allowed for the structure of this gene to be examined.

3.5.1 Characterisation of Basp1 sequence information

By comparing the protein sequences, it was clear that the entire protein coding region of the mouse *Basp1* gene was contained within the annotated exon. This sequence information was important for the design of the targeted knockout mouse model as we could be confident a full null phenotype would be generated by deleting this exon.

Two additional pieces of information, both important for the design of the mouse knockout model were gained from the initial RT-PCRs carried out on ES cell cDNA. The first being that *Basp1* was expressed in ES cells, allowing selection mechanisms to be used whilst targeting ES cells that rely on the expression of the gene of interest. The second being that there was a second exon of *Basp1*, 5' of the annotated exon. This combined information was used to design a mechanism of positive selection that allowed expression of a neomycin resistance gene from

the *Basp1* promoter whilst still deleting the protein coding region of the gene (see Chapter 4).

Since the updating of the sequence databases, the transcript size of the murine *Basp1* gene has been annotated as 1.7kb. This size correlates with the Northern hybridization size of ~2kb and the size of the cDNA sequence AK159486. It also closely matches the transcript size found by Northern hybridization in the cow (1.7kb) and rat (~1.5kb) (Park et al., 1998; Widmer and Caroni, 1990). A point to bear in mind was that this murine transcript size determined by Northern blotting was only accurate for *Basp1* transcripts that contain the probe sequences and are therefore able to form a hybridization product. The probe region was designed to the 5' region of the second exon and contains part of the protein coding region including the translational ATG start site, and was therefore assumed to contain all potential protein-coding transcripts of *Basp1*.

The size of the hybridization product from the Northern did not correlate with the transcript size estimated from the RT-PCR results. The second exon was found to extend past the 3' end of the cDNA sequences making the second exon ~4.8kb. Along with the first exon of *Basp1*, this predicts a transcript size of ~5kb. Although the cloned and sequenced 3' RACE products did not yield a match to the *Basp1* genomic region, the RT-PCR amplification products were strong and reproducible. There were three explanations for this anomaly. Firstly, there could be more than one *Basp1* spliceoform which were not all picked up by the Northern, either because they don't all contain the probe region, they are of low abundance or because the spliceoforms are tissue specific; the Northern was carried out with adult brain RNA whereas the RT-PCRs were carried out with ES cell RNA. Secondly, there could be another gene immediately 3' of the *Basp1* gene and the RT-PCR products were picking up this transcript. Thirdly, the RT-PCR products could be PCR artefacts. There is currently no evidence within the sequence databases for either of the first two possibilities.

The presence of only one hybridization product also suggested that there was no antisense transcript, at least not in the region overlapping the probe sequence which was around the transcriptional start site of *Basp1*. Since the novel

antisense transcript within the sequence databases was shown to overlap the transcriptional start site of *Basp1* and has subsequently been removed there is currently no evidence for the existence of a *Basp1* antisense transcript. It was also unlikely that the sense transcript size would be similar to an antisense transcript size thereby causing only one Northern hybridization product to be visualised, indeed, the antisense transcript was predicted to be only 531bp.

The rat *Basp1* gene has also been updated in the sequence databases. The two original exons that had overlapping sequence identity with the second exon of the human genome (see Figure 3.2) have been replaced with a single larger exon, and there is an additional upstream exon. The structure of the human, mouse and rat *Basp1* genes are now comparable.

3.5.2 *in situ* expression pattern

A probe was used for *in situ* hybridization experiments that contained the same region of the *Basp1* gene used for Northern blots i.e. the 5' region of the second exon containing part of the protein coding region including the translational ATG start site. As with the Northern blot result it was important to remember that any expression pattern seen was specific only for *Basp1* transcripts containing the probe sequence. By comparing the expression patterns of *Basp1* with the documented sites of *Wt1* expression at each developmental stage, it was possible to identify tissues of overlapping expression for both of these genes, suggesting the potential for a physiological interaction of the resulting protein products. Due to the identification of BASP1 as a component of the WT1 cosuppressor (Carpenter et al., 2004), a physiological interaction between Basp1 and Wt1 is expected. Although not all regions of documented *Wt1* expression have been examined for *Basp1* expression, most have been.

Where *in situ* hybridization has been carried out at the same developmental stages as RT-PCR the expression has correlated. Within the adult kidney however, expression was detected by RT-PCR but not by *in situ* hybridization.

E15.5

The newly identified sites of expression at this developmental stage were the choroid plexus arising from the fourth ventricle, trigeminal ganglion, thyroid gland, dorsal root ganglion, retina, intestines, testis, salivary glands and dermis. Expression within the central nervous system, lung and liver correlated with the BASP1 expression pattern determined by immunohistochemistry at the same developmental age (Carpenter et al., 2004). However, whereas Carpenter et al. (2004) saw BASP1 expression within the heart at E15.5, no *Basp1 in situ* expression could be seen within the heart at this developmental stage. Apart from the thymus and tongue which were not represented in the *in situ* sections, this expression study covered a wider range of organs and stages and has increased the knowledge of the *Basp1* expression pattern at E15.5.

Wt1 expression at E15-15.5 has been documented within the retina, the olfactory epithelium, the gonads, the mesorchium membrane that supports the testis and mesonephros, the kidney, testis, uterus, spinal cord and the roof of the fourth ventricle of the brain (Armstrong et al., 1993; Rackley et al., 1993; Wagner et al., 2002a; Wagner et al., 2005). Overlapping expression of *Basp1* and *Wt1* at this developmental stage therefore occurs in the retina, kidney, testis and spinal cord. Although there was *Basp1* expression within the choroid plexus arising from the fourth ventricle, there was no expression within the roof of the fourth ventricle.

E13.5

The newly identified sites of expression at this developmental stage were the cartilage of the jaw, tongue, olfactory epithelium and cartilage of the tail/limb. The expression pattern was very similar to that seen at E15.5. Although staining appeared more intense in certain tissues than at E15.5, *in situ* hybridization is not strictly a quantitative procedure and although the same probe and probe concentrations were used for each developmental stage, the *in situ* experiments were not carried out at the same time.

At E13 *Wt1* expression has been documented in the metanephric vesicles of the kidney and the sex cords of the gonads and in the olfactory epitheilum (Armstrong et al., 1993; Wagner et al., 2005). Overlapping expression of *Basp1* and *Wt1* at this developmental stage therefore occurs in the kidneys and olfactory epithelium.

E10-E10.5

The newly identified sites of expression at this developmental stage were the wall of the brain, forelimb bud, branchial arch, neural tube, the lining of the atrial chamber and the bulbus cordis. It was clear from the broad expression pattern at this developmental stage that *Basp1* was widely expressed early on in development and becomes progressively restricted as organs differentiate. This progressive restriction of *Basp1* expression has also been documented within chick development via immunofluorescence of tissue sections (Widmer and Caroni, 1990)

Wt1 expression has also been documented to become progressively restricted, from the first expression seen at E9 within lateral lining of the coelomic cavity and its surrounding epithelium to E19 when expression could only be seen in the kidney and gonads where it persists in the adult (Armstrong et al., 1993). At E10, *Wt1* expression could be seen in the mesothelial linings of organs contained within the coelomic cavity, the mesonephric vesicle and the olfactory epithelium (Armstrong et al., 1993; Wagner et al., 2005). Overlapping expression of *Basp1* and *Wt1* at this developmental stage therefore occurs in the mesothelial linings of at least some of the organs contained within the coelomic cavity.

Adult kidney

The *in situ* pattern did not show any specific expression of *Basp1* within the adult kidney, however, it was shown by RT-PCR that *Basp1* was expressed within the adult kidney (Figure 3.6). It has also been shown by immunohistochemistry with anti-BASP1 antibodies that Basp1 protein was expressed within the podocyte cells of the adult kidney (Carpenter et al., 2004). The *in situ* was repeated with the same result, and can not be put down to experimental procedure as an embryo section treated with the same probe at the same time did show staining. A positive probe known to pick up expression within the kidney was not used with these kidney samples, so lack of expression could be put down to a fixation problem or the RNA could have degraded prior to the start of the procedure. If organs remain in fixative for too long, it can inhibit the hybridization signal probably by inhibiting the ability of a probe to bind to the mRNA within a sample. Otherwise, the expression maybe too low to be detected by *in situ* hybridization but could be detected by the more sensitive procedure of RT-PCR.

Wt1 protein expression within the adult kidney has been documented within the podocyte cells by immunohistochemistry with anti-WT1 monoclonal antibodies (Carpenter et al., 2004). Interestingly, whereas human *Wt1* mRNA levels in the glomerulus were highest at foetal stages, human Wt1 protein levels in the glomerulus were highest in the adult, and mRNA could only be detected in the adult by RT-PCR (Mundlos et al., 1993).

Adult testis

Expression was seen within the seminiferous tubules in structures that were presumed to be the Sertoli cells. Wt1 protein has been shown to be expressed within the Sertoli cells of human testes by immunohistochemistry with anti-WT1 monoclonal antibodies (Mundlos et al., 1993), and *Wt1* mRNA has been shown to be expressed in rat and mouse Sertoli cells by *in situ* hybridization (Armstrong et al., 1993; Sharma et al., 1992). Expression of both genes within the same cells therefore suggests the potential for a physiological interaction.

Knocking out the Baspl gene in the mouse

4.1 Summary

In the previous chapter the characterisation of the *Baspl* expression pattern and genomic organisation within the mouse was discussed. The information gathered on the genomic organisation of this gene was essential to allow the design and generation of a *Basp*1 knockout mouse model. The generation of a *Baspl* knockout mouse model is the topic of this chapter. An introduction to the aims of this chapter is followed by a results section outlining the steps taken to construct a targeting construct that was used to then target murine embryonic stem cells *in vitro*. The processes carried out to verify the correctly targeted clones and suitability for *in vivo* use are also discussed. The results section continues with a summary of attempts to generate the knockout mice with these targeted embryonic stem cells followed by an analysis of these mice. A breeding knockout line was not generated as all *Baspl* heterozygous knockout (KO) mice born died postnatally. Attempts to understand the reasons for this postnatal lethality are then addressed.

4.2 Introduction

In order to elucidate the role Basp1 plays in development, particularly within the kidney, a mouse model was engineered lacking *Baspl*. In order to do this a gene targeting approach in embryonic stem cells (ES cells) followed by integration of these genetically modified cells into mouse embryos by the process of blastocyst injection was carried out (see section 1.2). Heterozygous and homozygous KO mice were to be monitored for signs of morbidity or mortality, initially by checking that the birth of animals of these genotypes occurs with the expected Mendelian frequencies. Morbidity studies to observe any early sign of kidney disease were to be carried out by monitoring the presence or absence of protein within the urine; the presence of protein being an early indicator of kidney malfunction (Menke et al., 2003). An analysis of the organs of these mutant mice was also to be carried out both prenatally and in adulthood to look for differences between those of wildtype siblings. Specifically, the weight, appearance and

histology of the organs that are known to express *Basp1* (see section 1.5.5 and Chapter 3) were to be examined. A more detailed study into the temporal and spatial expression pattern of *Basp1* was also to be carried out in these mice by virtue of the integrated β-galactosidase cassette. Due to the design of the targeting construct, a fusion product of the β-galactosidase gene with exon 1 of *Basp1*, should allow *lacZ* expression in amounts representative of the endogenous *Basp1* expression pattern, due to this cassette being driven by the intact *Basp1* promoter.

Any alteration in the levels of Wt1 was also to be observed in these mice. This could be carried out either by Northern or Western analysis of Wt1 in the organ lysates of null *Basp1* mice, or by RNA *in situ* hybridization or immunohistochemistry of Wt1 in sectioned organs of null *Basp1* mice. Additionally, these *Basp1* KO mice could be used to dissect the independent genetic roles Basp1 and Wt1 play during development by breeding these mice with a *Wt1* knockout line. *Wt1* KO mice were already available (Kreidberg et al., 1993) for this genetic breeding to take place.

Although a knockout of this gene in the mouse had already been generated and published by Frey et al. (2000), this was not the case when embarking on the project. The project was continued after this paper was published as it addressed neuronal brain function, specifically nerve sprouting and growth after muscle paralysis which was found to be lost in these knockout mice (see section 1.5.6), and not the potential role of Basp1 in kidney development and Wilms' tumour.

A systemic targeted knockout, rather than a conditional targeted knockout, was designed and generated. There were several reasons for this decision. With genes of unknown function it is advisable to design a full null in case the phenotype is very mild, as even when the sites of expression are known this is not a full indication that the gene expression is essential for the development or functional maintenance of this structure. In some cases another gene may show redundant expression, the paralogous *Hoxa11* and *Hoxd11* genes for example were shown to be functionally redundant during kidney development (Patterson et al., 2001); this was not assumed to be the case with *Basp1* as there were no known paralogs

in the mouse genome. The vector required for a conditional targeted knockout would also have been more technically demanding to construct (Nagy et al., 2003) and the resulting mice more expensive to generate as additional breeding steps and the maintenance of additional breeding stocks would have been required. Despite these disadvantages, the ability within a conditional targeted knockout to observe the effect of spatial and temporal elimination of a gene would be a major advantage, particularly if the gene was also required during development leading to death before the organ of interest could be investigated, as was the case with the *Wt1* systemic knockout (Kreidberg et al., 1993). In the event of embryonic lethality of our null mice it is still likely that the expression pattern of this gene and analysis of early organ development could be studied at these early embryonic stages or exclusively the phenotype of the heterozygous knockout mice could be studied.

Predictions can be made as to the possible kidney phenotype of *Basp1* KO mice by considering the role of Wt1 in kidney development by noting the evidence for Basp1 acting as a transcriptional cosuppressor through a direct interaction with Wt1 (Carpenter et al., 2004). The earliest urogenital expression of Wt1 in the mouse can be seen in the metanephric mesoderm cells at E9. These cells continue to express Wt1 and at E11 the epithelial ureteric bud arising from the Wolffian duct causes an inductive process whereby the mesodermal cells proliferate and condense to form the metanephric blastema around the ureteric bud and undergo a mesenchyme to epithelial transition. At E13 Wt1 is highly expressed in the differentiating epithelial cells of the mesonephros as they form the nephrons of the kidney (Armstrong et al., 1993). It is known that Wt1 is involved in the control of both apoptosis and proliferation during early kidney development. *Wt1* null mice show that this protein is required for the survival and early differentiation of the metanephric blastema as at E11 the ureteric bud failed to grow out of the Wolffian duct, the cells of the metanephric blastema underwent apoptosis and the inductive events required for the formation of the metanephros failed to occur (Kreidberg et al., 1993). In *Basp1* KO mice the absence of Basp1 protein will mean it will be unable to perform its cosuppressor role of binding to the suppression domain of Wt1. Therefore it is possible that the transcriptional activation domain of Wt1 will remain constitutively active with Wt1 target genes

being constitutively activated. In the case of *Wt1* null mice the ureteric bud is absent due to the failure of mesenchymal cells to differentiate; instead these mesenchymal cells degenerate via apoptosis highlighting the necessity of Wt1 for their survival. In this circumstance it seems that even without Wt1, apoptosis of kidney precursor cells is occurring. Maybe the opposite of this situation would happen when Basp1 is absent, with apoptosis being inhibited and overproliferation occurring. On the other hand it is possible that kidney development could not progress beyond a primitive stage without correct Wt1 control resulting in lack of kidneys as is the case with *Wt1* KO mice. Without further evidence for the role BASP1 and WT1 play as interacting proteins, this must remain a speculation.

It is apparent that BASP1 must also be acting independently of WT1 in some organs as it is expressed in regions that WT1 is not: particularly in the central nervous system where its expression pattern is diffuse. It is possible that *Basp1* KO mice could die of defects of the central nervous system, either developmentally or postnatally if, for example, the animals are unable to feed. What is more likely however considering that Basp1 protein is expressed in the heart during development (Carpenter et al., 2004) is that a defect in this organ could cause prenatal death or postnatal morbidity. It is known that *Wt1* KO mice die of heart failure; Wt1 being expressed in the mesothelial linings of the heart (Kreidberg et al., 1993).

4.3 Methods

Plasmids generated prior to the commencement of molecular cloning within this chapter can be found in Table 4.1. The sequences of two oligonucleotide primers used to generate a linker during one of the molecular cloning steps can be found in Table 4.2. Primer sequences used for PCR or RT-PCR within this chapter can be found in Table 4.3.

4.3.1 LacZ staining

The blue colouration is a result of a dark precipitation product formed when the enzyme beta-galactosidase expressed from the synthetic βgal reporter contained

Table 4.1

resource/plasmid name	reference/source	description
PAC library RPCI21	Osoegawa et al., 2000	source was female 129/SvevTACfBr mouse spleen genomic DNA
pMC1DTA	Yagi et al., 1990	based on pUC19 containing a DTA cassette driven by an MC1 promoter and terminated by SV40 DNA lacking a polyadenylation signal
P86a:1	unpublished. generated by Dr Andrew Ward	modified from pMC1DTA to include additional restriction sites for convenient sub-cloning of the DTA cassette
pSAβgeo	Friedrich and Soriano, 1991	based on pBluescriptII KS(-) containing a splice acceptor of the intron/exon2 adenovirus boundary, a fusion of the beta-galactosidase gene and neomycin resistance gene, and a polyadenylation signal from bovine growth hormone

Table 4.1. Details of the DNA elements used within the *Basp1* knockout cloning strategy that were generated prior to commencement of the project. Shows the reference or source of the PAC library and each plasmid and a brief description of the components that make up each resource.

Table 4.2

oligonuleotide primer pairs	sequences
CRESS1	$^{5'}$GGTGGCTGCAGGTCGACGC$^{3'}$
CRESS2	$^{5'}$GTCGACCTGCAGCCACCGC$^{3'}$

Table 4.2. Oligonucleotide primer pair used to form a linker during the *Basp1* knockout cloning strategy.

Table 4.3

primer pairs	sequences	Expected product size	annealing temp
BaspGenomicF	$^{5'}$TGCCATTTTGTGCTAAGGTGTTC$^{3'}$	600bp	60°C
BgalRS	$^{5'}$CATCTGCCAGTTTGAGGGGAC$^{3'}$		
BaspGenomicR	$^{5'}$CCTTCTCCAGAGTATCCGAC$^{3'}$	600bp	60°C
BgalRS	$^{5'}$CATCTGCCAGTTTGAGGGGAC$^{3'}$		
BaspGenomicF	$^{5'}$TGCCATTTTGTGCTAAGGTGTTC$^{3'}$	600bp	60°C
BgalRL	$^{5'}$AGTAACAACCCGTCGGATTCTC$^{3'}$		
3F	$^{5'}$GCTCCAACTCGCTCCACGCTC$^{3'}$	92bp	60°C
3R	$^{5'}$GGGTGGCTGGTGTCGTCCG$^{3'}$		
2F	$^{5'}$GGGCTACAATGTGAACGA$^{3'}$	255bp	60°C
2R	$^{5'}$CTCGCTCTTCTCGGGCTCGG$^{3'}$		
3F	$^{5'}$GCTCCAACTCGCTCCACGCTC$^{3'}$	429bp	60°C
2R	$^{5'}$CTCGCTCTTCTCGGGCTCGG$^{3'}$		
3Flong	$^{5'}$ACGCTCCAACTCGCTCCACGCT$^{3'}$	518bp	60°C
βgalRS	$^{5'}$CATCTGCCAGTTTGAGGGGAC$^{3'}$		
βgeoF2	$^{5'}$CCGAGCGAAAACGGTCTGCG$^{3'}$	565bp	60°C
βgeoR2	$^{5'}$CTTCCCGCTTCAGTGACAACG$^{3'}$		
3Fshort	$^{5'}$CTCCAACTCGCTCCACGCTC$^{3'}$	variable	60°C
βgalRL	$^{5'}$AGTAACAACCCGTCGGATTCTC$^{3'}$		
2FN	$^{5'}$GCAAAGCGTGCCGTCAAAGAG$^{3'}$	variable	60°C
βgalRL	$^{5'}$AGTAACAACCCGTCGGATTCTC$^{3'}$		
14F	$^{5'}$GAGTGCCATTGTATCCGTGAAC$^{3'}$	variable	60°C
βgalRL	$^{5'}$AGTAACAACCCGTCGGATTCTC$^{3'}$		

Table 4.3. Primer pairs used to perform PCR or RT-PCR. Shows primer names, sequences, the size product expected and the annealing temperature of the primer pairs.

within the targeting construct, breaks down X-gal (5'-bromo-4-chloro-3-indoyl-β-D-galactopyranoside).

4.4 Results

4.4.1 Targeting vector construction

Initially, the 129Sv strain RPCI mouse PAC library 21 (purchased from Human Genome Mapping Project, HGMP) (Osoegawa et al., 2000) was screened for the mouse *Basp1* sequence (work carried out by Dr Andrew Ward, University of Bath). The library was screened by radioactive hybridization to filters representing the available PACs using a mouse *Basp1* cDNA probe (IMAGE: 479433; donated by Dr Stefan Roberts, University of Manchester). PACs containing the required sequences were then digested with various restriction endonucleases and a Southern blot was carried out on these digested DNA fragments using the same mouse cDNA probe to identify genomic DNA fragments likely to contain the *Basp1* gene. Fragments of a suitable size for plasmid cloning were obtained with enzymes *HindIII*, *PstI* and *BglII*. PAC DNA digested with these enzymes was then sub-cloned into bacterial plasmids using a 'shotgun cloning' approach and colonies were screened for those containing the correct fragment by Grunstein colony hybridization (see section 2.1.14). Sequencing of the ends of the inserts within these sub-cloned plasmids was used to confirm the presence of genomic DNA from the *Basp1* gene and to identify which region of the gene had been inserted into each plasmid. These sub-cloned plasmids were digested with several different restriction enzymes in order to identifying unique restriction sites that could be used to produce a physical map, thereby providing enough molecular information for subsequent cloning of the targeting construct.

The targeting construct contained *Basp1* 3' and 5' homologous regions designed to completely displace the protein coding region of the gene contained within exon 2, a SAβgeo cassette and a diphtheria toxin-A gene cassette (DTA) (Nagy et al., 2003). The SAβgeo cassette contained a neomycin resistance gene (neo) that allows for positive selection within ES cells, fused to a *lacZ* reporter (βgal) to allow ready visualization of gene expression. The cassette also contains a

splice acceptor (SA) which allowed *Basp1* exon 1 to splice directly to this site forming a fusion transcript when expressed. The splice acceptor acts as an additional positive selection mechanism by allowing the promoter-less neomycin resistance gene to be expressed only if the construct integrates downstream of a transcriptional start site of a gene that is expressed in ES cells (as shown in Chapter 3, *Basp1* is expressed in ES cells). The DTA cassette allows for a further level of selection, the product of the diphtheria toxin-A gene being cytotoxic. This DTA cassette was placed at one end of the targeting vector allowing for negative selection within ES cells, as it should only be lost, and hence non-toxic, after correct genomic integration.

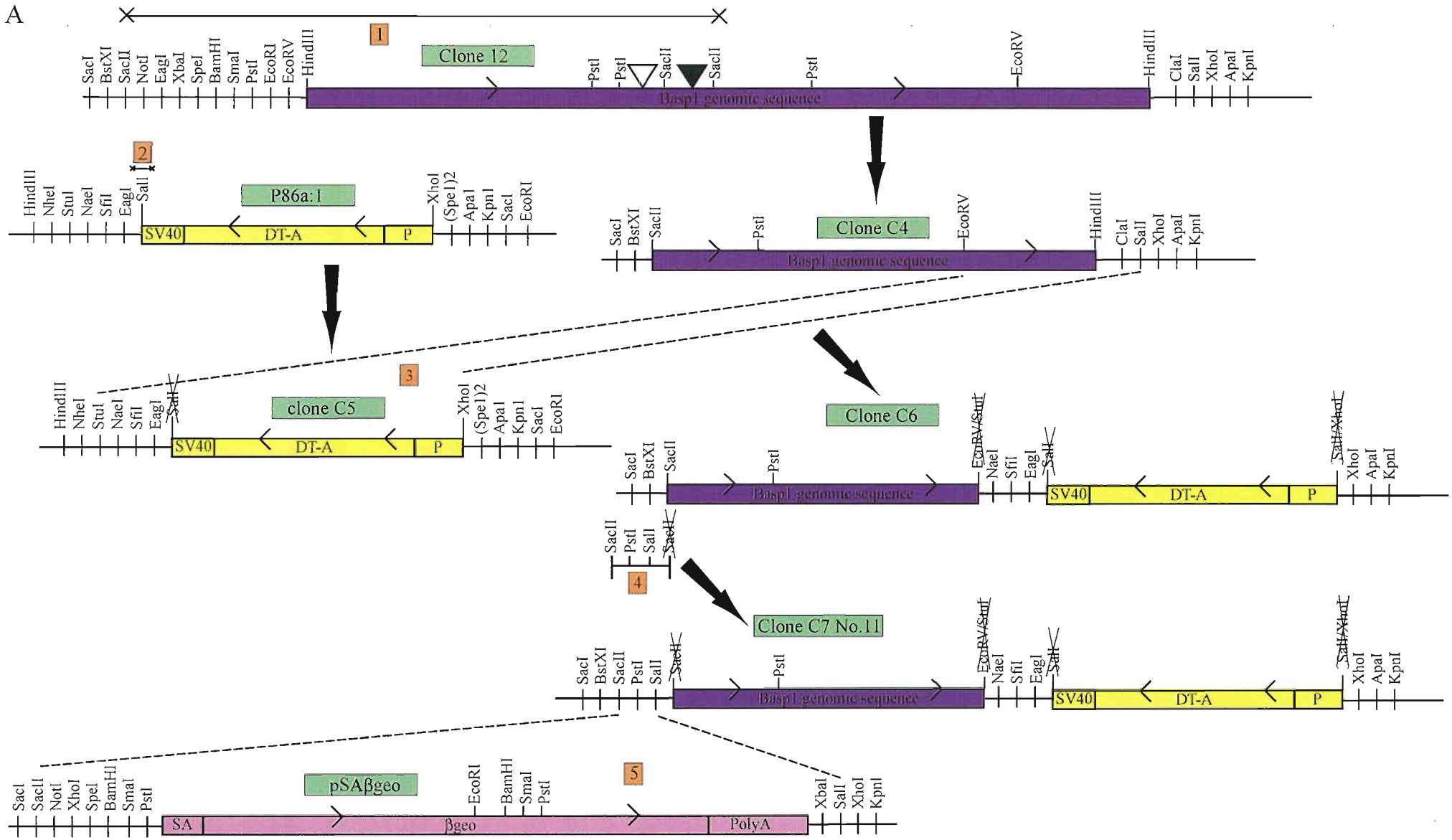
Figure 4.1A shows the six steps involved in creating the targeting construct. A detailed account of the cloning strategy and each sequential cloning step is discussed below. To check the correct fragments had been inserted into the relevant vectors at each step of the cloning strategy, plasmids were digested to look for expected restriction patterns, or PCR was performed using primers designed to amplify the region of interest. Following each cloning step, plasmids were sequenced to verify they contained the correct DNA fragments, as any mistakes would be built upon and therefore time consuming to rectify at a later date. Primer sequences can be found in Table 4.3.

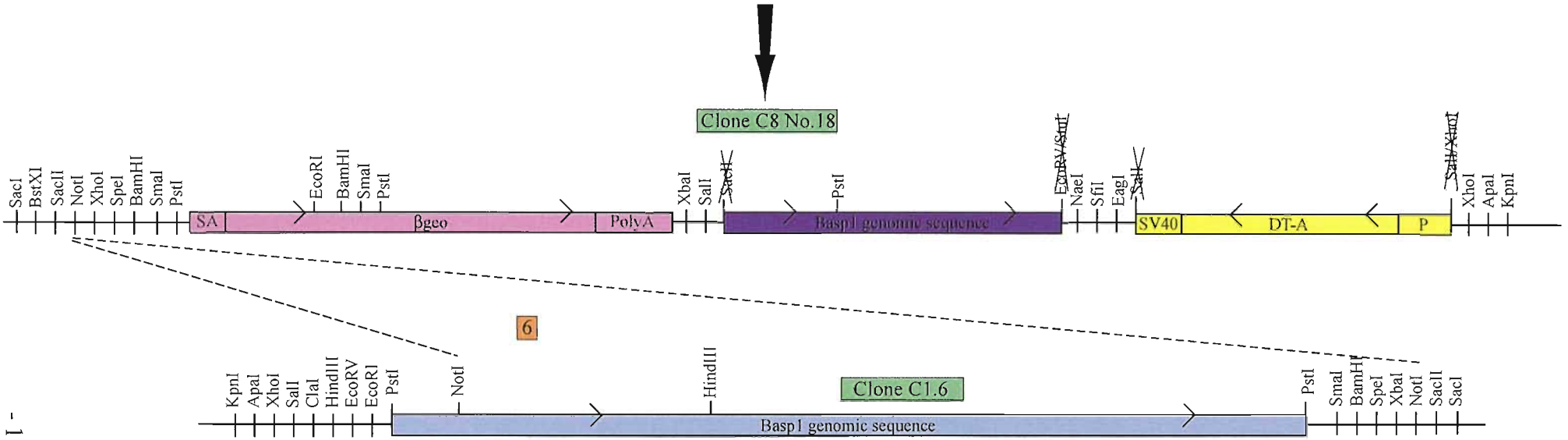
4.4.1.1 Cloning strategy

Step 1: A *HindIII* PAC fragment sub-cloned into pBluescript (named clone 12), containing part of the mouse *Basp1* gene, including the entire protein coding region, was used to generate the 3' homologous arm (Figure 4.1A). The region of *Basp1* genomic DNA used for this 3' homologous arm was approximately 3.4kb encompassing a *SacII* site at position 574bp within the single annotated exon of *Basp1* Ensemble v38 (entry ENSMUST00000058845), to an *EcoRV* site 3351bp downstream of this exon.

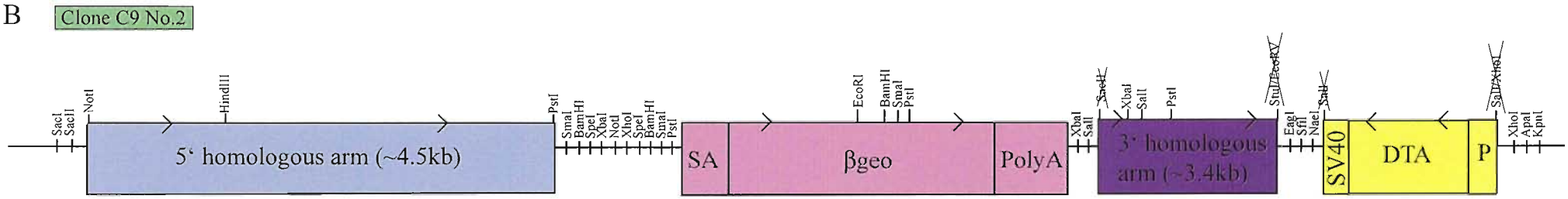
The *HindIII* plasmid, clone 12, was restriction digested with *SacII* giving rise to three fragments. The fragment containing the protein coding region of the *Basp1* gene between the *SacII* site in polylinker and the most distal *SacII* site within the

Figure 4.1. (A) The cloning strategy used to generate the *Basp1* targeting vector. Relevant restriction sites are shown within plasmid multiple cloning site regions (thin horizontal lines) and within plasmid inserts (contiguous coloured boxes). Numbered, orange boxes indicate sequential cloning steps. Labelled, green boxes indicate plasmid names. Black arrows point to the clones generated following each cloning step. 1. A *HindIII* PAC fragment sub-cloned into pBluescript (clone 12), containing part of the mouse *Basp1* gene, including the entire protein coding region (∇ start codon; ▼ stop codon) was used to generate the 3' homologous arm. The plasmid was restriction digested with *SacII* giving rise to three fragments. The fragment containing the protein coding region of the *Basp1* gene was removed by the re-ligation of the outermost *SacII* sites, minus the intervening region, to form clone C4. 2. As a unique *SalI* site was required at a later point in the construction, the *SalI* site in the DTA cassette (plasmid p86a:1) was removed to form clone C5. 3. The DTA cassette was then transferred into the 3' end of the 3' homologous arm. This was achieved by digesting the DTA cassette with the *SalI* site removed (clone C5) with *StuI* and *XhoI*, and the 3' homologous arm (clone C4) with *EcoRV* and *SalI*. *StuI* and *EcoRV* both generate blunt-ended cleavage sites that are therefore compatible for re-ligation. Similarly, *XhoI* and *SalI* generate compatible sticky-ended cleavage sites, with both ends destroyed when they are ligated together. The resulting construct was named clone C6. 4. A unique *SalI* site was needed in the 3' homologous arm in which to insert the SAßgeo cassette. A linker containing a *SalI* site with *SacII* overhangs at either end was inserted into the reformed *SacII* site of clone C6. Only one *SacII* site was re-created due to the design of the linker (this construct was named clone C7 N°11). 5. The SAβgeo cassette (from plasmid pSAβgeo) was then transferred upstream of the 3' homologous arm (clone C7 N°11) by digesting both constructs with *SacII* and *SalI*, and inserting the SAβgeo cassette into the *SacII* and *SalI* sites of the linker inserted in Step 4. It was important that the SAßgeo cassette was inserted in this orientation in relation to the direction of the *Basp1* gene, so that the splice acceptor (SA) can be spliced with the end of exon 1 of *Basp1* (this construct was named clone C8 N°18). 6. Finally, the 5' homologous arm was added to clone C8 N°18. A *PstI* PAC fragment sub-cloned into pBluescript (clone C1.6), containing part of the mouse *Basp1* gene, was used to generate the 5' homologous arm. Clone C1.6 was restriction digested with *NotI* giving rise to a fragment between the *NotI* site in the plasmid polylinker and a *NotI* site in the *Basp1* genomic fragment. The fragment containing part of the *Basp1* gene was then inserted into the *NotI* site of the SAβgeo cassette. (B) shows the completed construct named clone C9 N°2.


A
1
Clone 12
SacI BstXI SacII NotI EagI XbaI SpeI BamHI SmaI PstI EcoRI EcoRV HindIII
PstI PstI SacII SacII PstI EcoRV HindIII ClaI SalI XhoI ApaI KpnI
Basp1 genomic sequence
2
HindIII NheI StuI NaeI SfiI EagI SalI
P86a:1
XhoI (SpeI)2 ApaI KpnI SacI EcoRI
SV40 DT-A P
Clone C4
SacI BstXI SacII PstI EcoRV HindIII ClaI SalI XhoI ApaI KpnI
Basp1 genomic sequence
3
clone C5
HindIII NheI StuI NaeI SfiI EagI SalI
XhoI (SpeI)2 ApaI KpnI SacI EcoRI
SV40 DT-A P
Clone C6
SacI BstXI SacII PstI EcoRV/SmaI NaeI SfiI EagI SalI SalI/XhoI XhoI ApaI KpnI
Basp1 genomic sequence
SV40 DT-A P
4
SacII PstI SalI SacII
Clone C7 No.11
SacI BstXI SacII PstI SalI SacII PstI EcoRV/SmaI NaeI SfiI EagI SalI SalI/XhoI XhoI ApaI KpnI
Basp1 genomic sequence
SV40 DT-A P
5
pSAβgeo
SacI SacII NotI XhoI SpeI BamHI SmaI PstI EcoRI BamHI SmaI PstI XbaI SalI XhoI KpnI
SA βgeo PolyA


Clone C8 No.18
SacI
BstXI
SacII
NotI
XhoI
SpeI
BamHI
SmaI
PstI
SA
EcoRI
BamHI
SmaI
PstI
βgeo
PolyA
XbaI
SalI
PstI
Basp1 genomic sequence
NaeI
SfiI
EagI
SV40
DT-A
P
XhoI
ApaI
KpnI
6
KpnI
ApaI
XhoI
SalI
ClaI
HindIII
EcoRV
EcoRI
PstI
NotI
HindIII
Clone C1.6
Basp1 genomic sequence
PstI
SmaI
BamHI
SpeI
XbaI
NotI
SacII
SacI

B


Clone C9 No.2
SacI
SacII
NotI
HindIII
5' homologous arm (~4.5kb)
PstI
SmaI
BamHI
SpeI
XbaI
NotI
XhoI
SpeI
BamHI
SmaI
PstI
SA
EcoRI
BamHI
SmaI
PstI
βgeo
PolyA
XbaI
SalI
3' homologous arm (~3.4kb)
PstI
EagI
SfiI
NaeI
SV40
DTA
P
XhoI
ApaI
KpnI

Basp1 genomic fragment was removed by the re-ligation of the *SacII* sites minus the fragment containing this region. This remaining fragment sequence was to form the 3' homologous arm of the targeting vector. To check for plasmids containing the correct deletion, a *SacII* digest was carried out to look for a linearized fragment of ~7.5kb. This construct was named clone C4.

Step 2: As a unique *SalI* site was required at a later point in the construction, the *SalI* site in the DTA cassette (plasmid p86a:1) had to be removed. This site was removed by *SalI* digestion followed by filling-in of the 5' protruding ends of the *SalI* cleavage site and re-ligation of the subsequent blunt ends. To screen for recombinant plasmids a *SalI* digest was carried out to look for undigested plasmid DNA. This construct was named clone C5.

Step 3: The DTA cassette was then transferred from clone C5 into the 3' end of the 3' homologous arm. This was achieved by digesting the clone C5 with *StuI* and *XhoI*, and the 3' homologous arm (clone C4) with *EcoRV* and *SaI*. *StuI* and *EcoRV* are both blunt ended enzymes and were therefore compatible for re-ligation. *XhoI* and *SalI* ends can be directly ligated together, re-generating neither site. Note that destruction of the *SalI* site found in plasmid clone C4 was a requirement for this subsequent cloning event. To screen for recombinant plasmids, DNA was digested with *HindIII* and *XhoI* to look for a linearized fragment of ~7.5kb (the sequence between *EcoRV* and *SalI* of clone C4 containing a *HindIII* site should have been removed). This construct was named clone C6.

Step 4: A unique *SalI* site was needed in the 3' homologous arm in which to insert the SAβgeo cassette. A linker containing a *SalI* and a *PstI* site with *SacII* overhangs at either end was inserted into the reformed *SacII* site of the 3' homologous arm (clone C6). The two oligonucleotides used to form this linker were CRESS1/CRESS2 (Table 4.2). Only one *SacII* site was re-created due to the design of the linker. Recombinant plasmid DNA was digested with *SacII* alone and *SalI* alone to check for linearized fragments showing both sites were present. Recombinant plasmids were also sequenced to screen for the correct insertional orientation with the *SacII* site being 5' to the *SalI* site relative to the

orientation of the 3' homologous arm. The resulting construct was named clone C7 N°11 (or C6+linker N°11).

Step 5: The SAβgeo cassette (plasmid pSAβgeo) was then transferred upstream of the 3' homologous arm (clone C7 N°11) by digesting both constructs with *SacII* and *SalI*, and inserting the SAβgeo cassette into the *SacII* and *SalI* sites of the linker inserted in Step 4. It was important that the SAβgeo cassette was inserted in the same orientation as the *Basp1* gene, so that this cassette could be spliced with the end of exon 1 of *Basp1* by making use of the splice acceptor (SA). This is why the linker had to be in the correct orientation. Recombinant plasmids were screened by digestion with *SacII* and *SalI* to look for two fragments of ~7.5 and ~4.3kb. This construct was named clone C8 N°18 (or C7+SAβgeo N°18).

Additional techniques had to be carried out to achieve this cloning step, as it proved to be difficult. As it was not possible to assess whether both enzymes had cut clone C7 N°11 to completion (as the digested plasmid appeared as a linearized fragment of the same size with one or both enzymes), control ligations were carried out with vector DNA only to compare the number of colonies obtained from this transformation with that of vector plus insert; similar numbers of colonies were seen. The most obvious reason for these difficulties was that the digestion of the vector plasmid (clone C7 N°11) had not been sufficiently effective. A study carried out of the cleaving efficiency of several restriction enzymes close to the end of linearized DNA fragments (Moreira and Noren, 1995) showed that most enzymes required 6 base pairs on either side of their recognition site to cleave efficiently. With *SalI*, it was shown that this enzyme cleaved with 89% efficiency with only 3 base pairs on either side. *SalI* and *SacII* were 9 base pairs apart, a gap that should have been adequate, yet it is possible that the efficiency was not 100%. Grunstein colony hybridization (see section 4.1.14) was carried out with a βgeo specific probe to try to detect colonies containing the correct insert. Only one colony appeared to be specifically detected by the probe. Colonies around this area of the agar plate were grown up in liquid culture and plasmid DNA was extracted. None of these colonies contained the correct insert. Therefore, this technique was abandoned and the

cloning steps were repeated making sure that optimum buffers were used for restriction digests, and plasmid DNA from a large number of colonies were checked (76 in total) before one with the correct insert was found. The bacteria containing this plasmid were difficult to grow both on agar plates and in liquid culture (possibly due to the large size of the insert; now 11.8kb). The plasmid was transferred into XL10-Gold® ultracompetent cells (Stratagene) (see section 2.1.6) which solved the problem. This might have been the reason it was difficult to find colonies containing the correct insert during the earlier attempts at this cloning step.

Step 6: Finally, the 5' homologous arm was added to clone C8 N°18. A *PstI* sub-cloned PAC fragment in pBluescript (named clone C1.6), containing part of the mouse *Basp1* gene, was used to generate the 5' homologous arm. The region of *Basp1* genomic DNA used for this 5' homologous region was approximately 4.5kb, encompassing a *PstI* site 199bp upstream of the single annotated exon of *Basp1* Ensemble v38 (entry ENSMUST00000058845) to approximately 4700bp upstream of the single annotated exon, where the plasmid was cut at a *NotI* site. Restriction mapping of the *Basp1* PAC sub-clones placed this site 5' of the *HindIII* site. However, although the *HindIII* site was later identified within the *Basp1* genomic sequence, the *NotI* site was not and we therefore assume that the *NotI* site is polymorphic between the mouse strains or has been acquired from the PAC vector during sub-cloning of clone C1.6.

Clone C1.6 was restriction digested with *NotI* giving rise to two fragments between a *NotI* site in the plasmid polylinker and a *NotI* site in the *Basp1* genomic fragment. To aid separation of the insert (~4.5kb) from the backbone (~3.5kb), the plasmid DNA was also digested with *PvuI* that only cuts the backbone of pBluescript, thereby separating this fragment into smaller pieces. The fragment containing part of the *Basp1* gene was then inserted into the *NotI* site of the SAβgeo cassette. Ligation products were transformed into XL10-Gold® ultracompetent cells. Recombinant plasmid DNA was screened by digestion with *NotI* to look for two fragments of ~11.8kb and ~4.5kb. Since the 5' homologous arm needed to be in the same orientation relative to the 3' homologous arm and the SAβgeo cassette, plasmids containing the insert were

sent for sequencing. Sequencing failed for all plasmids, probably due to the low molar concentration of DNA associated with a plasmid of such a large size (~16.3kb). A PCR protocol was therefore designed to screen for plasmids containing the insert in the correct orientation. Primers were designed to amplify a product between the 5' homologous arm in either direction and the SAβgeo cassette [BaspGenomicF/βgalRS or BaspGenomicR/βgalRS]. Diluted plasmid DNA was used as the template for this PCR. A PCR product of ~600bp could only be amplified with primers BaspGenomicF/βgalRS if the insert was in the correct orientation. Conversely, a PCR product of the same size could only be amplified with primers BaspGenomicR/βgalRS if the insert was in the wrong orientation. This final ~16.3kb construct was named clone C9 N°.2 (or C8+C1.6 N°.2) (see section 4.6 for details of how to find this plasmid within the laboratory).

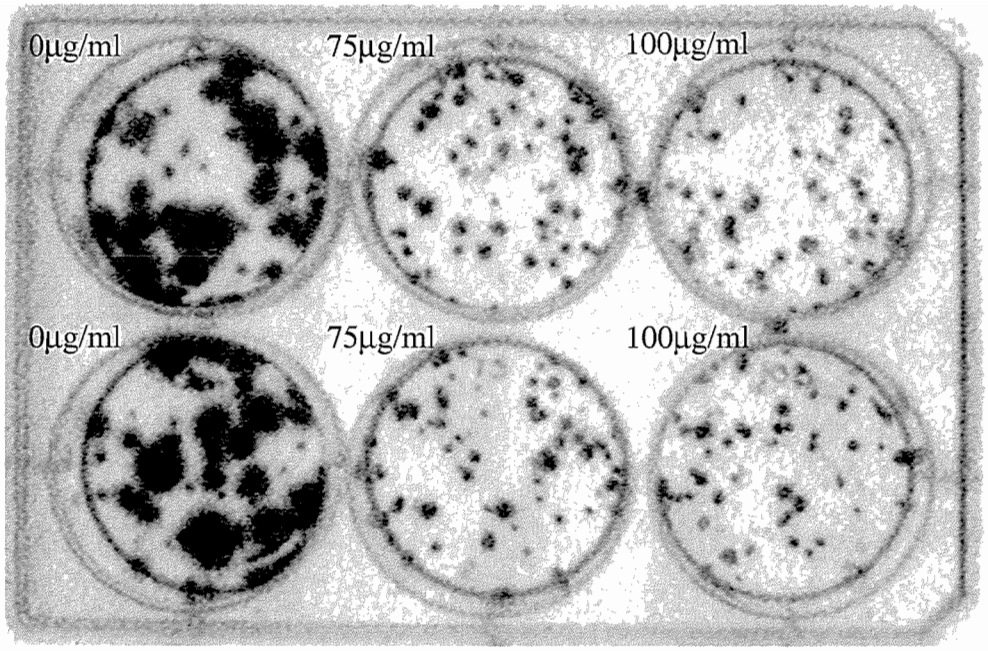
Once cloning was completed, the final construct (C9 N°.2) (Figure 4.1B) was linearized at the 5' end, using the unique *SacII* restriction site, and introduced into ES cells derived from the 129P2 strain of mice by electroporation.

4.4.2 Introduction of the targeting construct into ES cells

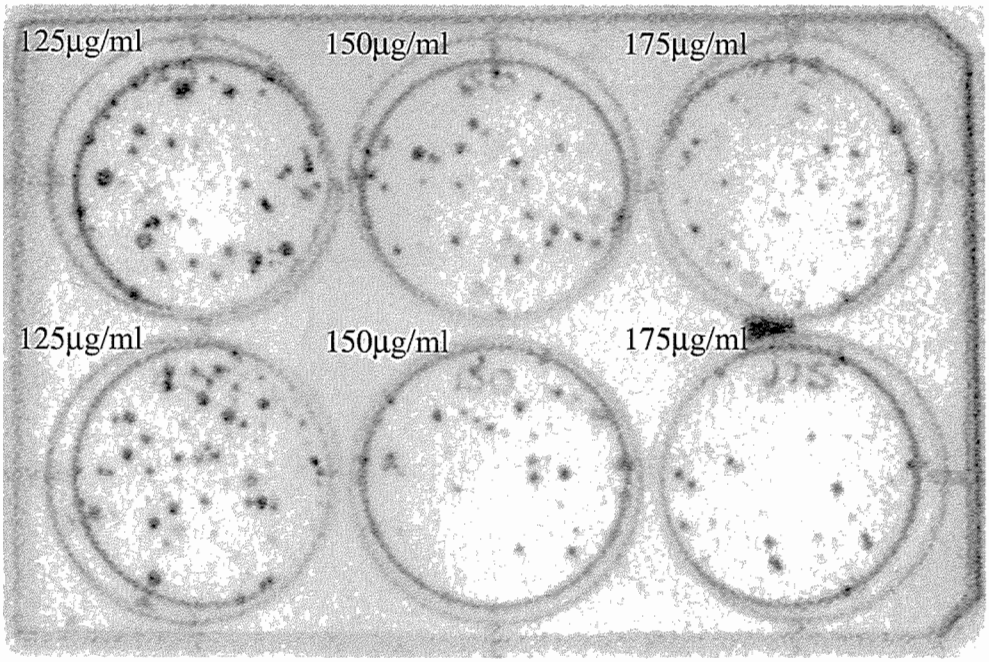
Transfection of the *Basp1* targeting construct into embryonic stem cells was carried out using electroporation. E14C4 ES cells (Smith, 1991) were used (see section 2.2.1). Initially it was necessary to carry out an antibiotic dose response test to ascertain the concentration of Geneticin® G418 sulphate required to kill the parental geneticin susceptible ES cell line. This diagnostic test (detailed in section 2.2.5) involved subjecting ES cells to varying G418 concentrations over a period of 14 days, followed by staining with methylene blue to detect living cells (Figure 4.2). The minimum G418 concentration required to kill all non-transfected, geneticin susceptible E14C4 ES cells was determined to be 225μg/ml. This concentration of G418 was used in transfection experiments. Electroporation was carried out as described in section 2.2.6, with a single electrical pulse of 300V, 250μF. These settings were found to be optimum for DNA uptake and integration into the E14C4 ES cells used, as determined by control transfections with constructs containing either green fluorescent protein

Figure 4.2. Tissue culture dishes containing E14C4 ES cells stained with methylene blue following 14 days of treatment with G418 antibiotic (each concentration is duplicated). Black regions are comprised of living cells stained with methylene blue. As can be seen in the first 6 well dish (A), many cells were still alive with G418 concentrations ranging from 0-100μg/ml. Less cells survived over 14 days with G418 concentrations ranging from 125-175μg/ml as shown in the second dish (B). The wells corresponding to 200μg/ml G418 in the third dish (C) have a couple of remaining living cells, however there are no remaining living cells with ≥225-250μg/ml G418. Therefore the lowest dose of Geneticin® G418 sulphate that could be used to kill all parental E14C4 ES cells grown in the presence of LIF, 15% FCS and without feeder cells was 225μg/ml.

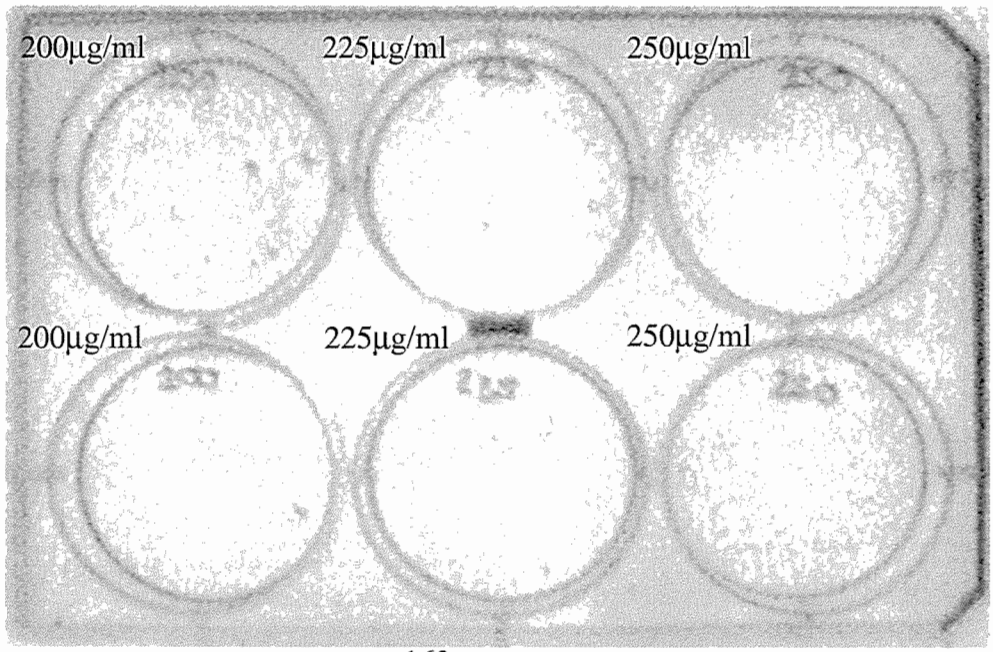
A


0μg/ml
75μg/ml
100μg/ml
0μg/ml
75μg/ml
100μg/ml

B


125μg/ml
150μg/ml
175μg/ml
125μg/ml
150μg/ml
175μg/ml

C


200μg/ml
225μg/ml
250μg/ml
200μg/ml
225μg/ml
250μg/ml

(GFP) under the control of the CMV promoter that can be visualised when exposed to UV light or the *lacZ* gene under the expression of the Simian vacuolating virus 40 (SV40) promoter that upon expression forms a blue colouration product in the presence of X-gal. Control transfections were carried out in tandem as described in section 2.2.6 but varying the electrical pulse conditions. 9µg of the SV40-βgal and CMV-GFP constructs were added to a fixed density of ES cells, then electroporated under the following conditions: 240V, 500µF; 300V, 250µF; 800V, 3µF (these are common electroporation conditions used with mammalian cells, (Nagy et al., 2003)). After electroporation, $^{1}/_{6}^{th}$ and a $^{1}/_{12}^{th}$ volume of ES cells were plated into gelatin-coated tissue culture petri dishes (10cm diameter), and allowed to recover for one day before staining with X-gal or looking for green fluorescing cells under UV light. There were a comparable number of blue cells with the 240V, 500µF and 300V, 250µF conditions, but no blue cells with 800V, 3µF conditions. There were few green fluorescing cells compared to *lacZ* positive cells, therefore only the SV40-βgal construct was used to further home in on the optimum transfection conditions. Next, 9µg of the SV40-βgal construct was electroporated under the conditions laid out in Table 4.4, and blue stained cells in 10 fields of view were counted. The SV40-βgal construct was ~6.4kb; 9µg was used during electroporation as this amount of DNA was roughly equivalent to 25µg of the 16.3kb targeting construct. Apart from 310V, 250µF where the cells were killed, all other test conditions showed similar numbers of blue cells and therefore have similar transfection efficiencies. As 300V, 250µF gave a higher average count, these were the parameters chosen for transfection of the targeting construct.

25µg of *SacII* linearized targeting construct C9N°2 was transfected with E14C4 ES cells and grown in geneticin selective media as described in section 2.2.6. The passage number (P) of the parental E14C4 ES cell line was P18/P19.

4.4.3 Identification of correctly targeted ES cell clones

Following transfection of ~8 x 10^7 ES cells with 25µg of *SacII* linearized targeting construct, 36 geneticin resistant clones were recovered and analysed by PCR and Southern hybridization. To verify integration of the targeting construct

Table 4.4

Electroporation conditions	Field of view colony counts										Average counts
230V, 500μF	2	5	1	1	1	4	2	1	2	2	2.1
240V, 500μF	4	2	2	5	1	3	2	1	2	3	2.5
250V, 500μF	0	8	1	0	1	2	2	7	1	2	2.4
290V, 250μF	1	2	1	0	2	2	2	2	5	2	1.9
300V, 250μF	2	4	1	4	3	5	5	3	4	2	3.3
310V, 250μF	0	0	0	0	0	0	0	0	0	0	0

Table 4.4. Numbers of blue stained cell clones counted per field of view following treatment with X-gal after being transfected with a βgal construct under differing electroporation conditions. Each stained colony represents an original single cell that had taken up the SV40-βgal construct and is therefore expressing beta-galactosidase. The more stained cell clones seen per field of view, the more effective the transfection. 300V and 250μF are shown to be the most effective transfection conditions.

by PCR, genomic DNA was amplified using primers designed to the 5' homologous arm and the SAβgeo cassette [BaspGenomicF/βgalRL]. This PCR should give a product of approximately 600bp if the construct (or at least the 5' most region of the construct) has integrated into the cell line (Figure 4.3A). Due to the size of the homologous arms (~4.5kb and ~3.4kb) it was not possible to amplify across this distance to verify if the construct had been targeted in the correct genomic location, therefore, Southern hybridization was carried out.

For Southern hybridization, genomic DNA was digested with *EcoRI*, and probes were designed 5' and 3' of the homologous arm sequences (Figure 4.4A-C) to hybridize to both an endogenous ~13kb wildtype allele and a smaller targeted allele. Using the 5' probe (probe C1.6), the band size expected for the targeted allele was 8.75kb, and for the 3' probe (probe C4) was 7.85kb. Of the 36 surviving clones, 30 clones (83%) appeared to be correctly targeted (Figure 4.4D-F). To verify that only targeting of one genomic location had occurred, the ES cell clone DNA was also hybridized to a probe designed to identify the SAβgeo cassette. The band size expected for a single targeting event was 8.75kb when using this SAβgeo probe. A band of this size was seen with the majority of samples (Figure 4.4G), suggesting only one integration site of the SAβgeo cassette within these clones. All geneticin resistant clones were checked via Southern hybridization. Only clones used for blastocyst injection were verified using PCR.

4.4.4 Identification of suitable clones for generating transgenic mice

Before extensive time and resources were placed in generating transgenic mice using one of the correctly targeted ES cell clones, further tests of suitability were carried out so that the optimum clones could be chosen.

An *in vitro* method to check that the *Basp1* gene had been knocked-out and that the splice acceptor was working as it should in targeted clones was to check that an RNA *Basp1*:βgeo fusion transcript of this gene was made. As exon 1 of *Basp1* should be spliced onto the splice acceptor of the SAβgeo cassette during transcription of the targeted allele of *Basp1*, primers were designed to amplify a

Figure 4.3. Screening ES cell clones for homologous recombination at the *Basp1* locus by PCR and RT-PCR. (A) Agarose gel showing PCR products obtained with primers BaspGenomicF/βgalRL (primers in the 5' homologous arm and the SAβgeo cassette of the targeting construct), which should amplify a product only if the targeting construct has been integrated into the cell line. Clones Basp3 and Basp4 are targeted cell lines and show the expected band (BaspGenomicF/βgalRL 601bp). (B) Agarose gel showing RT-PCR products obtained with primers 3F/2R (primers in exon 1 and exon 2 of *Basp1* respectively), which should pick up the wildtype *Basp1* transcript; and primers 3Flong/βgalR (primers in exon 1 of *Basp1* and the SAβgeo cassette of the targeting construct), which should pick up a transcript only if the targeting construct has integrated downstream of the *Basp1* transcriptional start and is spliced with exon 1. E14C4 is the untargeted wildtype cell line and shows only a wildtype product. Clones Basp3 and Basp4 are correctly targeted cell lines and show both the wildtype product (3F/2R 429bp) and the spliced *Basp1*:βgeo fusion product (3Flong/βgalRS 518bp). Lanes represent RT-PCR reactions carried out with first strand synthesis products made with (RT +) or without (RT -) reverse transcriptase, or a water control (0).

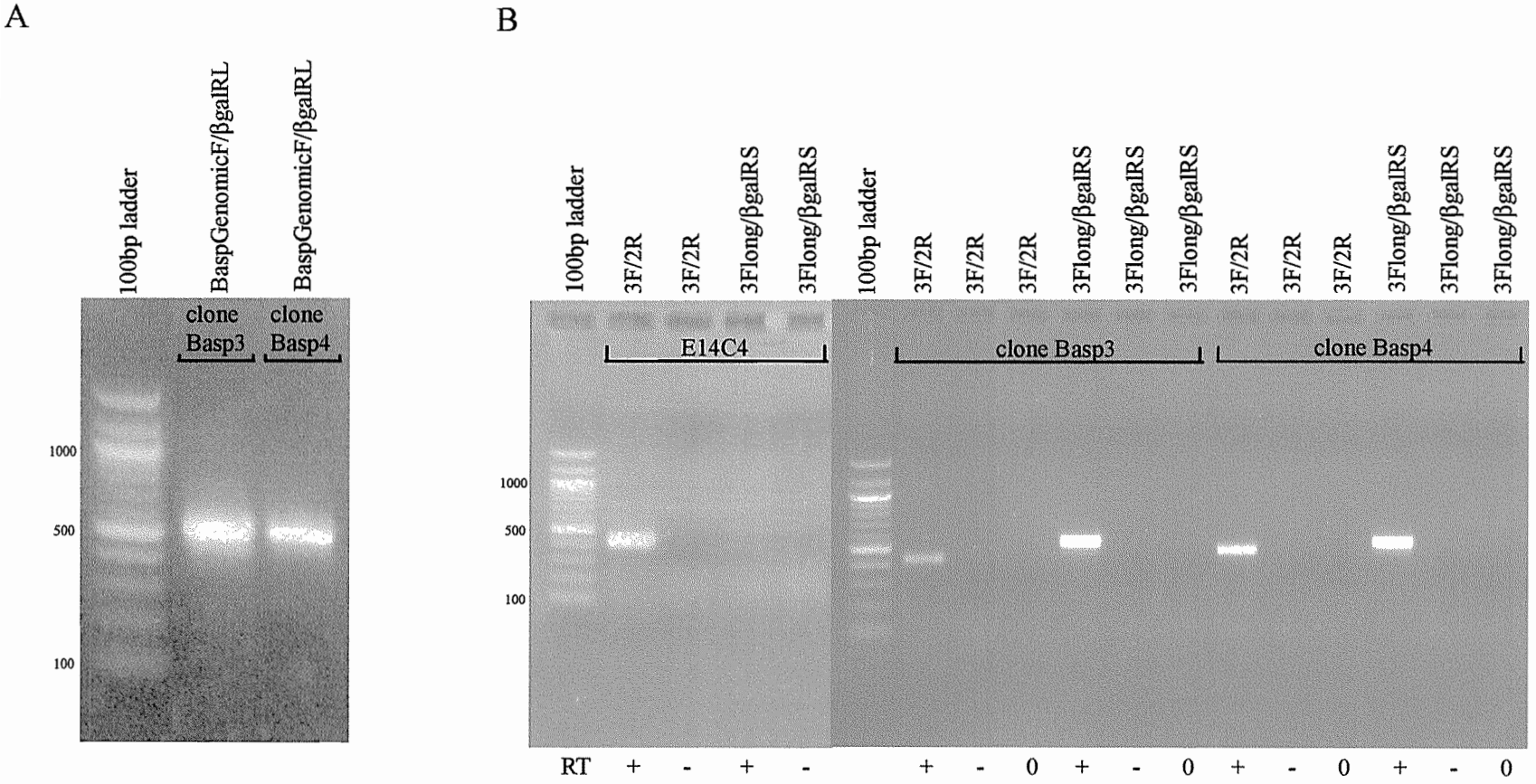
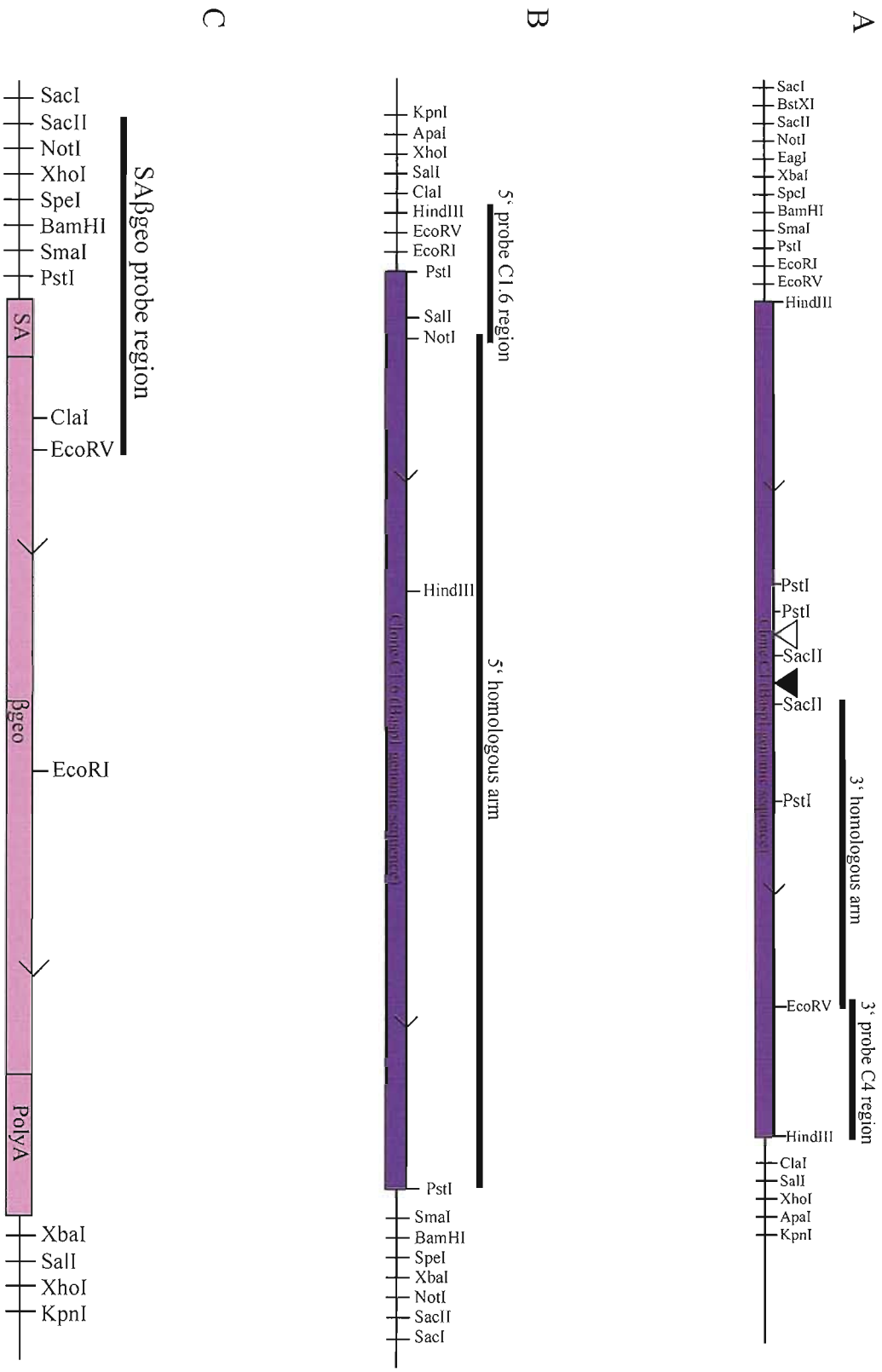
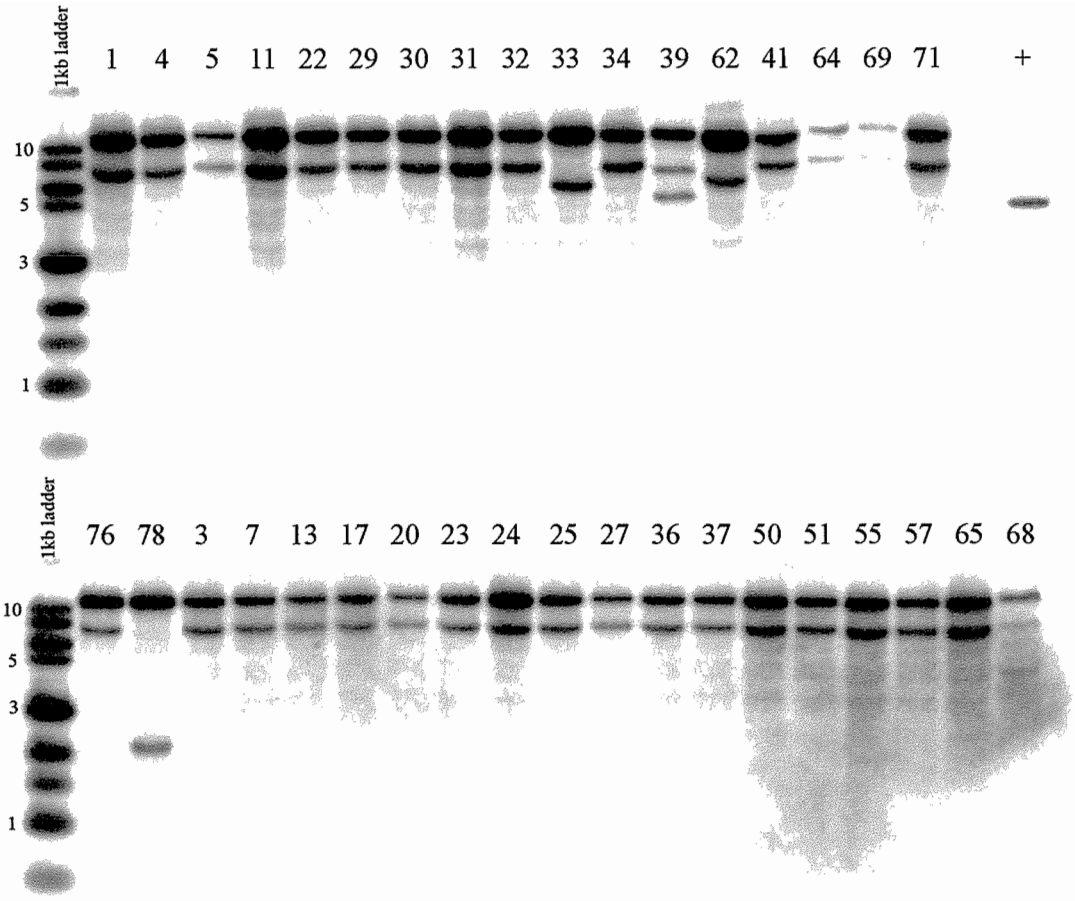

A
100bp ladder
BaspGenomicF/βgalRL
BaspGenomicF/βgalRL
clone Basp3
clone Basp4
1000
500
100
B
100bp ladder
3F/2R
3F/2R
3Flong/βgalRS
3Flong/βgalRS
E14C4
clone Basp3
clone Basp4
1000
500
100
RT + - + - + - 0 + - 0 + - 0 + - 0

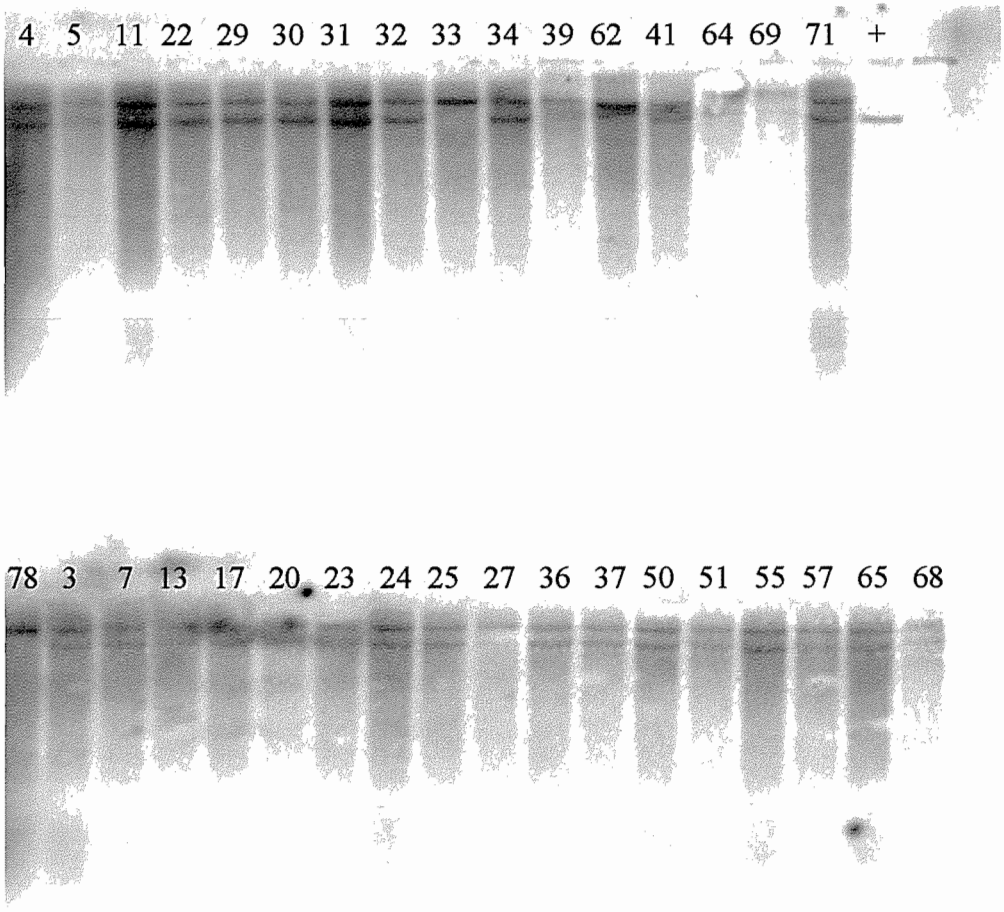
Figure 4.4. Probes used for Southern blotting and the autoradiograph results of the Southern blotting experiments. A-C, Schematic representations of the position of (A) 3' probe C4, (B) 5' probe C1.6 and (C) SAβgeo probe in relation to the homologous arms or SAβgeo cassette. D-F, Southern blotting results showing DNA hybridisation banding patterns representing neomycin resistant ES cell clones with (D) 3' probe C4, (E) 5' probe C1.6 and (G) SAβgeo probe. The upper bands of each sample in blots D and E represent the ~13kb wildtype band, and the lower bands represent the targeted size. D; The expected size of 7.85kb for correctly targeted cells was present in 31 of the samples (samples Basp33, Basp39, Basp62, Basp78, and Basp68 appeared to be incorrectly targeted). E; The expected size of 8.75kb for correctly targeted cells was present in 29 of the samples (samples Basp27, Basp33, Basp39, Basp62, and Basp78 were incorrectly targeted, 2 samples were missing). The (+) lanes contain 10pg of plasmid DNA used as positive controls for the probes (plasmid clone C4 in D; plasmid clone C1.6 in E). F; shows a comparison between the banding patterns of three samples probed with the 5' and 3' probes. G; The expected size of 8.75kb with probe SAβgeo representing a single targeting event was present in 28 of the samples (Basp39, Basp62, and Basp78 were incorrectly targeted, 5 samples were either missing or undetectable).


A
SacI
BstXI
SacII
NotI
EagI
XbaI
SpeI
BamHI
SmaI
PstI
EcoRI
EcoRV
HindIII
PstI
PstI
SacII
SacII
PstI
EcoRV
HindIII
ClaI
SalI
XhoI
ApaI
KpnI
3' homologous arm
3' probe C4 region
B
KpnI
ApaI
XhoI
SalI
ClaI
HindIII
EcoRV
EcoRI
PstI
SalI
NotI
HindIII
PstI
SmaI
BamHI
SpeI
XbaI
NotI
SacII
SacI
5' probe C1.6 region
5' homologous arm
C
SacI
SacII
NotI
XhoI
SpeI
BamHI
SmaI
PstI
SA
ClaI
EcoRV
βgeo
EcoRI
PolyA
XbaI
SalI
XhoI
KpnI
SAβgeo probe region

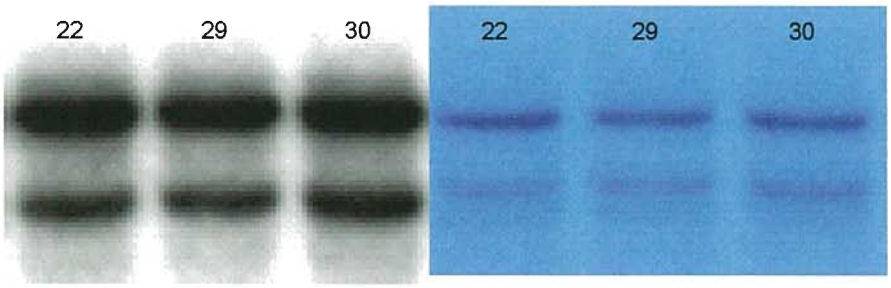
D


1kb ladder
1 4 5 11 22 29 30 31 32 33 34 39 62 41 64 69 71 +
10
5
3
1
1kb ladder
76 78 3 7 13 17 20 23 24 25 27 36 37 50 51 55 57 65 68
10
5
3
1

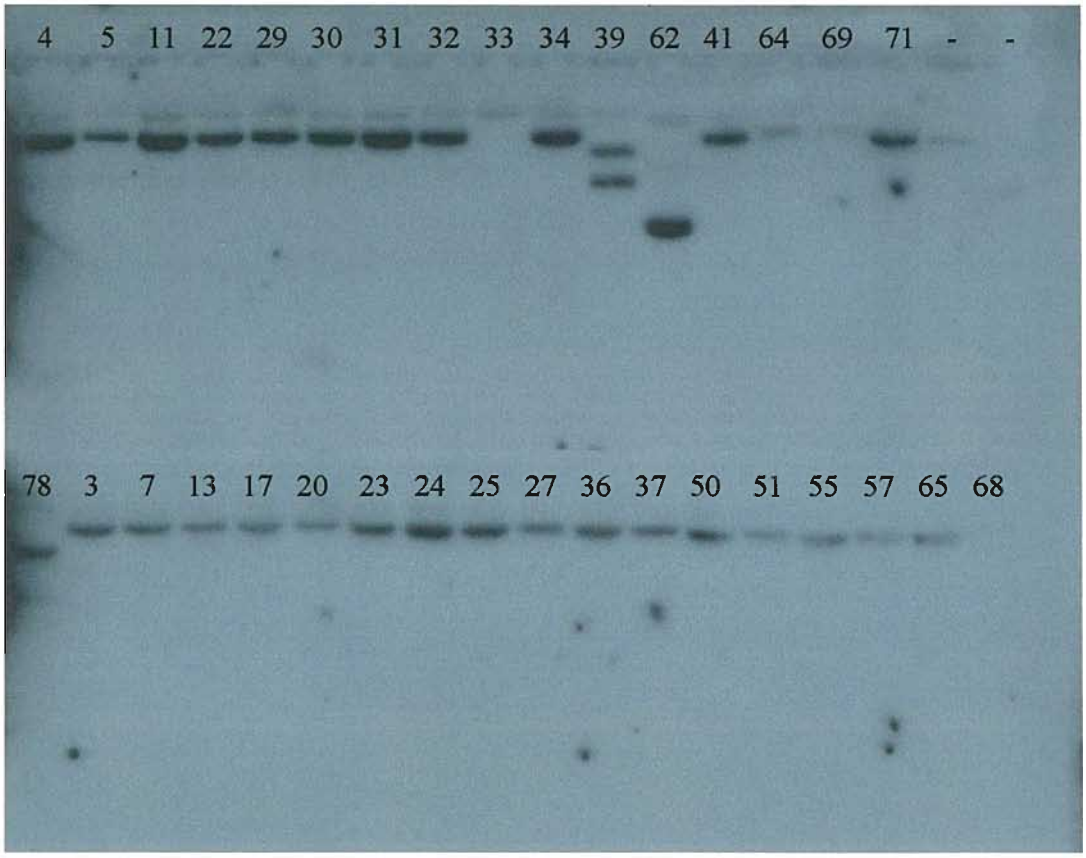
E


4 5 11 22 29 30 31 32 33 34 39 62 41 64 69 71 +
78 3 7 13 17 20 23 24 25 27 36 37 50 51 55 57 65 68

F


22
29
30
22
29
30

G

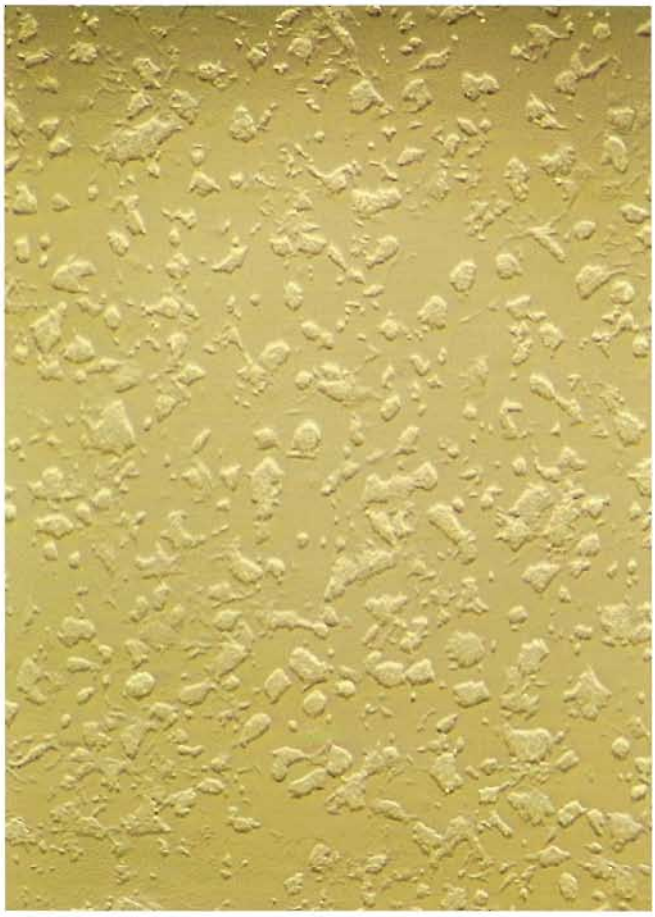

4 5 11 22 29 30 31 32 33 34 39 62 41 64 69 71 - -
78 3 7 13 17 20 23 24 25 27 36 37 50 51 55 57 65 68

product between the first exon of *Basp1* and the SAβgeo cassette [3Flong/βgalRS]. In the non-targeted allele, exon 1 of *Basp1* should be spliced onto the splice acceptor of exon 2 and control primers were therefore designed to amplify between these two exons [3F/2R]. Total RNA was extracted from targeted ES cell clones, cDNA was produced, and RT-PCR was carried out with primers 3F/2R and 3Flong/βgalRS. The product sizes expected were 429bp and 518bp respectively (Figure 4.3B). (Only clones used for blastocyst injection were verified using RT-PCR).

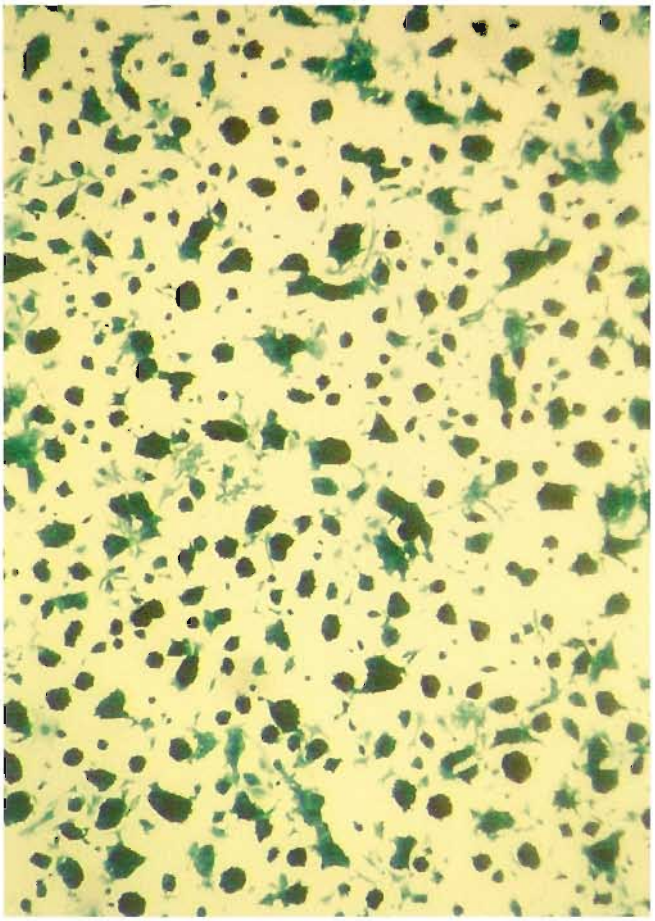
It was also important to verify that the β-galactosidase reporter was working correctly, and to check this, ES cells were stained for *lacZ* expression (see section 2.2.8). Cells that had correctly integrated the SAβgeo cassette showed a blue colouration product upon staining with X-gal, whereas untargeted cell lines, or incorrectly targeted cells did not (Figure 4.5). All ES cell clones used were positive for *lacZ* expression.

ES cells were also observed for correct morphology, making sure that the majority of clones within a population had formed colonies of highly refractive cells with smooth boundaries, as is characteristic of undifferentiated embryonic cell. Chromosomes spreads of some clones were prepared to check that the correct number of chromosomes remain after genome manipulation. It is known that a dramatic reduction in the likelihood of germline transmission occurs in clones in which <75% of metaphases are euploid (Longo et al., 1997) and chromosome counts were therefore considered to be a potentially valuable time saving procedure. Chromosomes were visualised using DAPI and counted to establish how many metaphase spreads contained the correct chromosome number (40 in the mouse). As this procedure in itself was laborious, only two cell lines were initially chosen for this analysis. Targeted ES cell lines Basp55 and Basp65 were the two chosen. Cell line Basp55 exhibited 73% euploidy and Basp65 exhibited 57% euploidy (Figure 4.6). Although these two cell lines exhibited <75% euploidy, both were used for blastocyst injections, cell line Basp55 having the higher percentage of euploid metaphase spreads being used a greater number of times compared to cell line Basp65.

Figure 4.5. Photographs of ES cells treated with X-gal. (A) Parental ES cell line E14C4 showing no blue staining in the presence of X-gal. (B) Correctly targeted ES cell line Basp65 showing blue staining in the presence of X-gal. The βgal reporter is linked to and driven by the *Basp1* promoter, showing that this gene is expressed in ES cells.



A



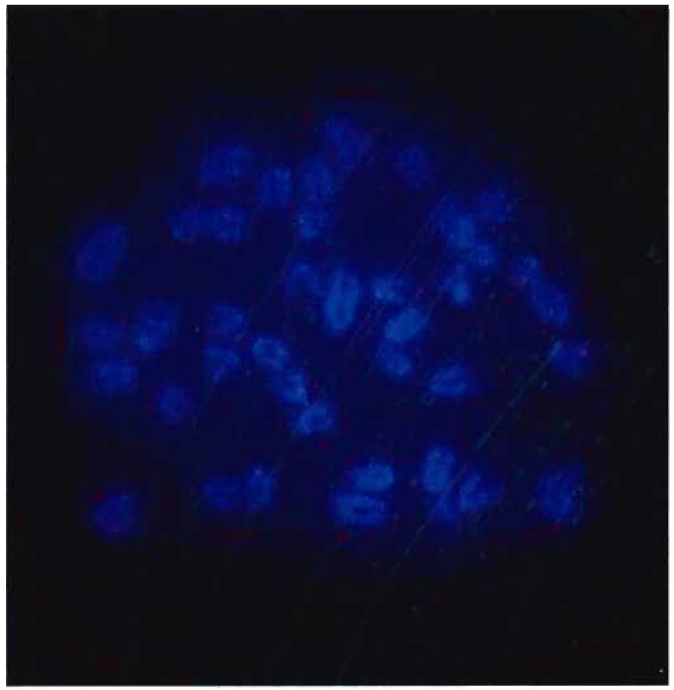
B

Figure 4.6. (A) The numbers of chromosomes counted in ten spreads for the targeted ES cell lines Basp55 and Basp65. (B-C) Chromosome spreads of targeted ES cells lines. (B) Basp55 and (C) Basp65 stained with DAPI in order to ascertain karyotype counts. Spreads were visualized using fluorescence microscopy with ultraviolet light at 600X magnification with an oil immersion lens.

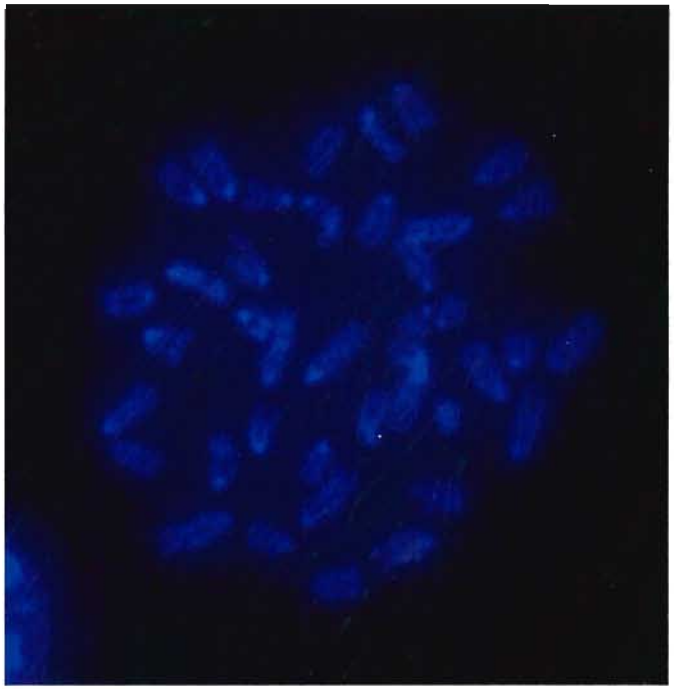
A

Chromosome Counts											
Basp55	40	40	40	40	39	40	39	41	40	40	40
Basp65	41	39	40	40	41	40	40	-	-	-	-

B



C



4.4.5 Blastocyst Injections

ES cell lines Basp55 and Basp65 were the two cell lines initially used for blastocyst injections, these two cell lines satisfying several tests for suitability as discussed in the previous section.

These cell lines were used repeatedly (Table 4.5), however, when the chimeras generated from these two cell lines failed to give germline transmission, a different strategy was employed, aimed at testing empirically a greater number of ES cell clones. Clones were used for injections in which tests of suitability where satisfied (excluding time consuming chromosome spreads), making sure to carry out only a couple of injection sessions with each cell line. Table 4.5 summarises all of the blastocyst injection trials and also shows the efficiency of blastocyst injections. As can be seen by analysing the numbers involved, higher numbers of injected blastocysts did not always lead to higher numbers of live offspring. Likewise, higher numbers of live offspring did not always result in a high percentage of chimeras. Only one ES cell clone (Basp3) gave rise to a single chimera that exhibited germline transmission of the *Basp1* KO allele.

4.4.6 Chimeric mice exhibiting morbid features

It was noted that four mice generated from two distinct targeted cell lines (Basp55 and Basp65) exhibited morbid features and had to be culled. One of these mice was chimeric by coat colour whereas the other three mice were only potentially chimeric as they did not appear to be chimeric by coat colour or PCR. The number of morbid mice typically seen in wildtype breeding colonies is minimal, the reasons for these illnesses therefore needed to be addressed.

These mice had a range of problems including abnormal motor behaviour - stumbling and fitting; small body; the appearance of a squashed face. One explanation for these observations is that these mice were physically damaged during the process of blastocyst injection. Alternative explanations could be that the resultant phenotype of having chimeric cells with only one copy of *Basp1* is very severe or that additional mutations had been introduced into the ES cells that were used to generate these chimeric mice and this was causing the severe

Table 4.5

ES cell line/ Parameter	Basp1	Basp3	Basp4	Basp23	Basp30	Basp51	Basp55	Basp65
No. transfers	3	2	5	3	1	3	12	7
No. blastocysts transferred	41	25	57	32	17	44	238	73
No. live births	0	6	11	0	5	5	45	14
% transferred blastocysts resulting in live births	0	24	19	0	29	11	19	19
No. chimeric offspring	0	1	2	0	3	2	11	0
% live births resulting in chimeras	0	17	18	0	60	40	58	0
No. chimeras generating founder lines	0	1	0	0	0	0	0	0
Overall % transferred blastocysts resulting in live births	**16**							
Overall % live births resulting in chimeras	**22**							
Overall % of chimeras generating founder lines	**5**							

Table 4.5. Efficiency of blastocyst injections to generate *Basp1* knockout founders. Shows the efficiency of obtaining live births, chimeras and founders from blastocyst injections for each targeted cell line used, as well as overall efficiencies.

phenotypes observed. These latter explanations are unlikely as other chimeras made from the same cell lines were viable, although the extent of chimerism or the specific tissues colonised by these mutant ES cells could have played a role.

4.4.7 Breeding Chimeras

If some of the injected ES cells contribute to the gonads they will give rise to gametes of the ES cell-derived genotype.These gametes may be sperm or oocytes, but they are usually sperm; this is because most ES cell lines in use are XY in origin and this genotype causes a sex conversion of the chimeric embryos during development to become male (the E14Tg2A-clone4 ES cells that have been used were XY in origin). If these ES cells are heterozygous for the targeted mutation (as they were in our case), then 50% of the chinchilla offspring that arise from gametes of these ES cells should be heterozygous for the targeted mutation and 50% should be wildtype, as they have inherited a wildtype allele from both parents.

Each chimeric animal was bred with a wildtype mouse of the inbred 129 strain. Coat colour observations could therefore be used to ascertain whether germline transmission had taken place. This is because the ES cells used originated from 129 strain mice that have a recessive chinchilla coat, meaning that only pups showing a chinchilla coat need to be screened further (agouti pups have no ES cell-derived genes and can be culled). An ear clip was taken from chinchilla pups so that DNA could be extracted and the mice genotyped with primers specific for the SAβgeo cassette [βgeoF2/βgeoR2]. Chimeric animals that didn't give rise to any chinchilla pups were culled after producing approximately 50-100 pups, unless they had stopped breeding, had died, or had been used for *in vitro* fertilisation (IVF) before reaching this number (Table 4.6).

4.4.8 Attempts to establish a *Basp1* knockout breeding line by *in vitro* fertilisation

To practice *in vitro* fertilisation (IVF) (see section 2.4.6), F1 hybrid mice were used as gamete donors, as hybrid strains are optimum for this process (F1 mice

Table 4.6

Chimera Name	No°. of offspring born	Germline transmission?	No°. of KO+/- offspring born	% of KO+/- offspring
Badger ♂	24	×	0	0
Brainy ♂	74	×	0	0
Butch ♂	105	×	0	0
Basil ♂	44	×	0	0
Bonzo ♂	12	×	0	0
Bonkers ♂	17	×	0	0
Brutus ♂	315	✓	12	3.8%
Bramble ♀	52	×	0	0
Basher ♂	0	×	0	0
Babe ♀	49	×	0	0
Belle ♀	0	×	0	0
Bart ♂	11	×	0	0
Beans ♂	9	×	0	0
Bruno ♂	50	×	0	0
Bracken ♀	52	×	0	0
Boxer ♂	34	×	0	0
Busker ♂	63	×	0	0

Table 4.6. Breeding record of *Basp1* chimeras and the generation of a single *Basp1* knockout founder. Shows numbers of pups born to each chimera, indicating those that included ES cell-derived genes and the percentage that were heterozygous *Basp1* knockout founders in the case of germ line transmission.

used in this study were always the first generation of a cross between strains C57BL/6 and CBA) (Vergara et al., 1997). Out of 49 oocytes, 38 were successfully fertilised (78% fertilisation rate), as seen by the progression to the two-cell stage. As this was a trial experimental procedure these embryos were not re-implanted.

Two male chimeras (Brainy and Bart) were used for IVF with the aim of increasing the potential number of founder animals. Brainy was a chimera that had stopped breeding, whereas Bart was a younger male that was still breeding. F1 females were used as oocyte donors with sperm collected from the male chimeras. Out of approximately two hundred oocytes, only 12 were fertilised (<6% fertilisation rate). These 12 embryos were re-implanted and 2 pups were born, neither of which were founder animals. Inbred animals are documented to be poor at fertilisation *in vitro* (Vergara et al., 1997), and as the chimeric animals would have had a mixture of C57BL/6 and 129 sperm (both inbred stains), the possibility of success was considerably lower than with F1 hybrid animals.

4.4.9 A *Basp1* knockout chimera exhibiting germline transmission

The single chimera that transmitted the ES cell-derived genes including the *Basp1* KO allele through the germline was named Brutus. In a text book example of germline transmission from chimeric mice, 50% of the offspring that arise from gametes of ES cells with heterozygous deletions should be heterozygous for the targeted mutation and 50% should be wildtype. In the case of Brutus this predicted Mendelian inheritance pattern did not occur. Instead, of the 129 chinchilla pups born from this chimera crossed with a 129 female (agouti pups were discounted as they do not contain ES cell material), all were found to have wildtype genotypes (i.e. only inherited the wildtype ES cell allele). This problem became apparent when the absence of the *Basp1* KO allele could no longer be readily put down to chance. Possible reasons for these observations included: that the heterozygous phenotype was lethal *in utero*, or that a second mutation linked to the first, either spontaneous in nature or caused during the culturing and manipulation steps, was present in the genome and causing neonatal death. A third possibility was that there could have been a strain specific lethality for the

mutation caused by modifier genes (genes that have small quantitative effects on the level of expression of another gene). Finally, it was possible that the ES cells used to generate the chimera were of mixed origin, i.e. even though the cells were all geneticin resistant, some cells could have had the targeted mutation and others may not have. In this circumstance, a 50:50 ratio of untargeted:targeted pups within the 129 offspring would not be expected.

After these problems were highlighted it was decided that Brutus should be bred to females of the F1 outbred strain. These females are better breeders in that they give birth to larger litters and take better care of their young. This provided a different strain to show if this factor was influencing the survival rates and also helped us in obtaining larger numbers of pups in case a spontaneous mutation was causing lethality. If a second linked mutation was to blame for lethality, it would be possible to find a pup without this spontaneous mutation but with the targeted mutation due to recombination. Of course, the closer two loci are linked the larger the cohort of offspring needed to find this combination of alleles. No heterozygous knockout mice of this strain were born within the first few litters.

The possibility of a heterozygous lethal phenotype was addressed by dissecting pregnant females (Brutus x F1 or 129 wildtype) at different stages of gestation. A tail biopsy was taken from each pup followed by fixation and staining in X-gal staining solution overnight to detect the presence of β-galactosidase activity (see section 2.4.7), indicative of the presence and expression of the *lacZ* cassette at the targeted allele. The *lacZ* gene codes for an enzyme called beta-galactosidase that can cleave X-gal. Upon oxidation of one of the resultant by-products, an insoluble blue product is formed. Since *lacZ* is under the control of the *Basp1* promoter, this colouration acts as an expression reporter for this gene.

E12.5 pups were dissected from an F1 female. In the presence of X-gal, out of 11 pups dissected, 1 went blue, and was also positive when genotyped by PCR. This shows that these mice can survive up till E12.5, but doesn't answer if they are being conceived in a Mendelian ratio and dying before birth. In order to answer this second question a 129 female was used to distinguish by eye pigmentation whether pups were derived from the ES cells. E15.5 embryos were dissected

from a 129 female and a tail biopsy and *lacZ* staining was carried out as before. Out of 10 pups dissected, 7 went blue and were positive for the *Basp1* KO allele when genotyped (Figure 4.7). Also, all 10 pups had a lack of pigmentation in their eyes characteristic of the recessive phenotype of the 129 strain of mouse and therefore must all have had an ES cell contribution. Therefore, these mice were still surviving up until E15.5, and within the bounds of the expected Mendelian ratio. Since Brutus' litter sizes up until this point were not obviously smaller than expected, these heterozygous knockouts might have been the first to have been conceived, especially if the ES cell contribution to the germline was of mixed origin.

Brutus was therefore allowed to continue to breed in order to ascertain whether any live heterozygous knockouts would be born. Brutus had 27 litters (totalling 154 offspring) before any heterozygous knockouts were born, however none of the heterozygous knockouts that were born survived to breeding age (see section 4.4.10).

4.4.10 Phenotypic analysis of *Basp1* heterozygous knockout mice

Of the 12 *Basp1*$^{+/-}$ pups that survived past birth, five were found dead in cage at 11, 19, 22, 25 and 46 days after birth and seven had to be culled at 18, 19, 22, 25 (three mice), and 27 days after birth due to the appearance of morbid features (hunched over, lack of grooming, hair on end), a failure to thrive (very thin) and an awkward gait (walking on toes and stumbling). Both males and females were observed with these characteristics. The genotype of these mice was determined by an ear clip PCR with primers specific for the SAβgeo cassette [βgeoF2/βgeoR2]. In some cases, body weight data was recorded and weight differences between heterozygous knockout mice and wildtype siblings were seen. For example, in one litter a heterozygous mouse in week three following birth weighed 6.9g whereas a wildtype littermate of the same sex weighed 10.5g. In a separate litter another heterozygous mouse in week three following birth weighed 5.2g whereas two of its wildtype litters mates of the same sex weighed 15.3g and 16.2g.

Figure 4.7. Photograph of an E15.5 Basp+/- embryo showing blue X-gal staining in positions of *Basp1* expression. There is strong expression in the brain, tongue, vertebrae, lungs and umbilical cord. There does not appear to be any staining in the heart or liver. Other organs are not visible in this whole mount.

Five of the culled mice were dissected to look for obvious anatomical abnormalities and organs from four of these mice were X-gal stained to look for the presence of β-galactosidase indicative of the localisation of *Basp1* expression. The kidneys and testes of one of these mice were sectioned and stained to look for histological abnormalities (Figures 4.8 and 4.9). Kidneys were stained with PAS and testes with H&E (see sections 2.3.5 and 2.3.4). Apart from their small body size and therefore proportionally small organ sizes compared to wildtype littermates, there were no obvious abnormalities internally, except the stomach and intestines were filled with air (not liquid). There was also no evidence of protein being present in the urine (tested via a coomassie stained protein gel (see section 2.1.18), an early indication of kidney disease (Menke et al., 2003). The first three mice culled and dissected had their whole organs stained for *lacZ* expression. *lacZ* staining was seen in the spleen, pancreas, lymph nodes, thymus, testis, ovaries, oviduct, uterus, bronchial vessels of the lungs and regions of the brain (Figure 4.10). Strong *lacZ* expression could be seen in the cerebrum, median eminence and pituitary gland. Weak or no expression was present in the cerebellum, pineal gland, optic nerve, and medulla oblongata (the olfactory bulbs were not dissected from the mouse. Staining was also seen in the kidneys, stomach, intestines and faintly in the liver, but this staining was also present in the wildtype control, therefore if any specific β-galactosidase expression was present in these organs it had been masked by endogenous staining. There was no expression in the heart (both externally and internally). Due to the high endogenous *lacZ* staining in certain organs, and the fact that whole organs stained for β-galactosidase did not retain this staining when embedded in wax and sectioned, when the fourth mouse was culled and dissected another approach was taken to try to get around these limitations. This approach involved obtaining cryosections of the organs and staining these sections on slides, (see section 2.4.8) however this method did not eliminate the endogenous staining in wildtype organs.

Histology of the kidneys and testes did not reveal anything unexpected: development of both organs was delayed in comparison to wildtype, but the main structural features were present (Figures 4.8 and 4.9). Within the kidney S-shaped bodies were still detectable although along with mature glomeruli. The

Figure 4.8. (A-D) PAS staining of kidney sections from wildtype and heterozygous Brutus offspring (Brutus is a Basp1 knockout chimera). Magenta staining shows the presence of glycoproteins and blue staining shows nuclei. A, wildtype kidney (40x); B, heterozygous kidney (40x); C, wildtype kidney (200x); D, heterozygous kidney (200x). c, cortex; m, medulla; d, distal convoluted tubule; p, proximal convoluted tubule; s, s-shaped body; g, glomerulus; a, interlobular artery.

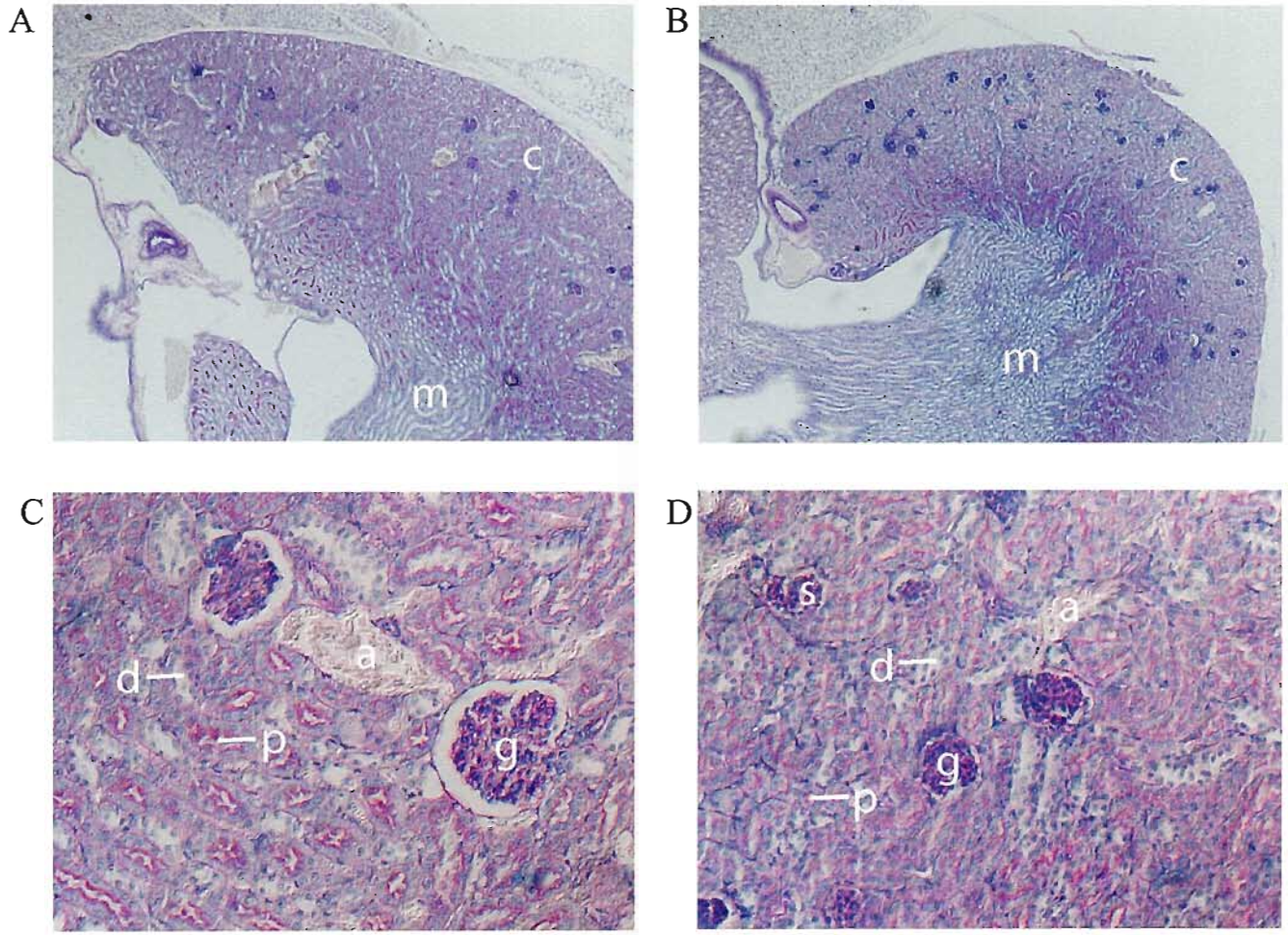

A
c
m
B
c
m
C
d
a
p
g
D
s
a
d
g
p

Figure 4.9. (A-D) Haematoxylin and eosin staining of testis sections from wildtype and heterozygous Brutus offspring (Brutus is a Basp1 knockout chimera). Blue staining shows nuclei and purple staining shows cytoplasm. A, wildtype testis (40x); B, heterozygous testis (40x); C, wildtype testis (100x); D, heterozygous testis (100x). s, seminiferous tubule; l, leydig cells.

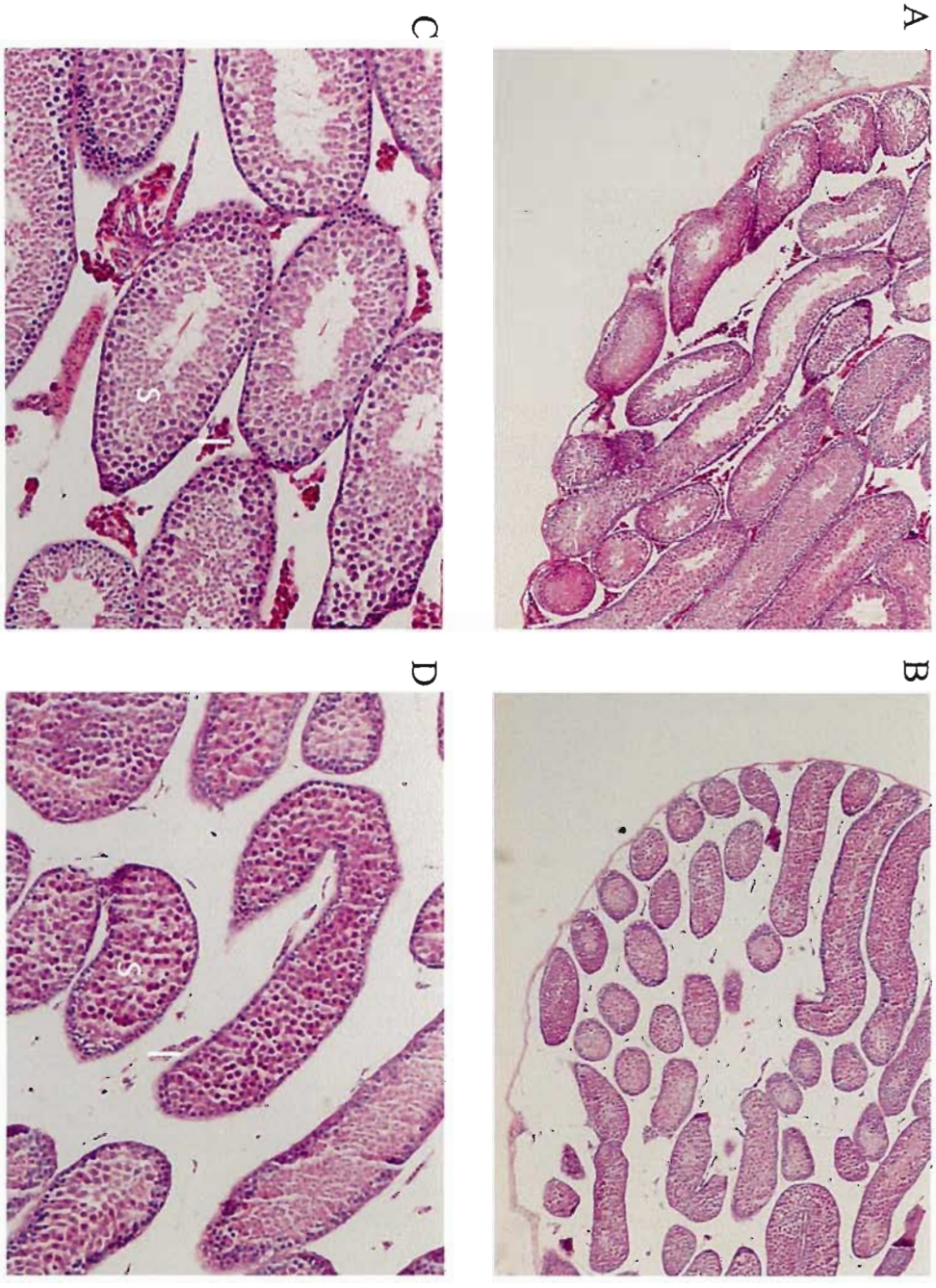
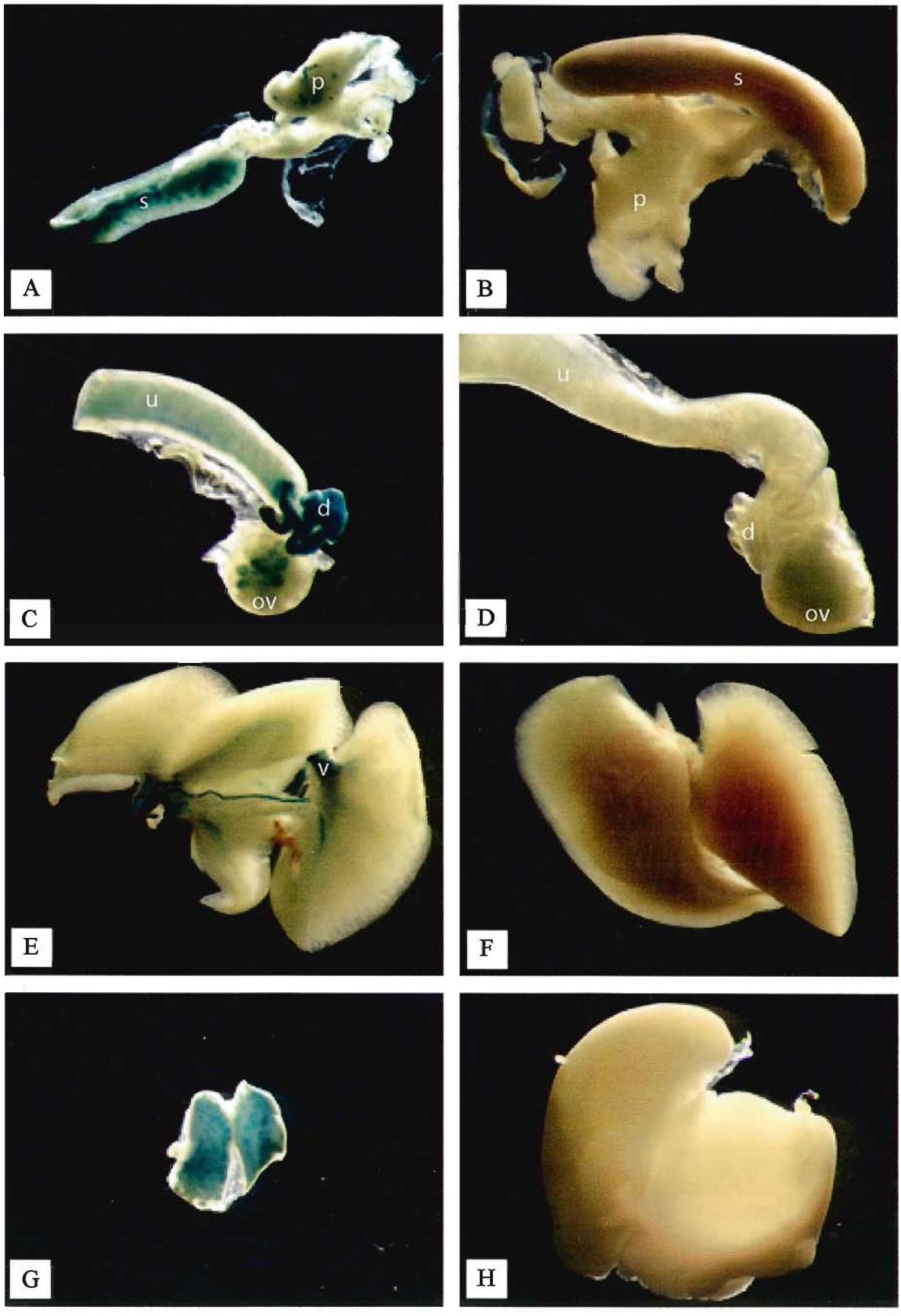
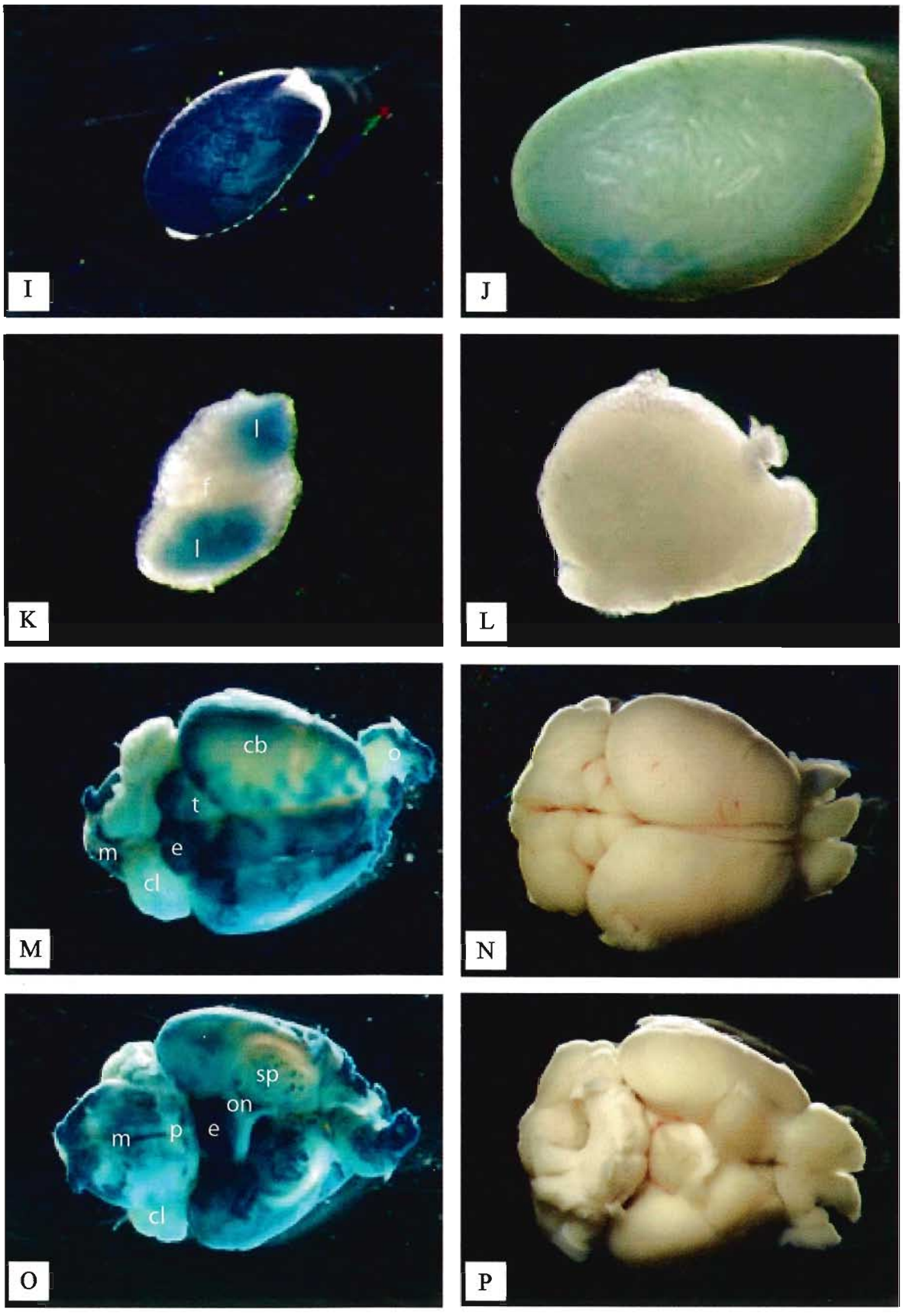

A
B
C
D

Figure 4.10. (A-P) *LacZ* staining of whole organs from heterozygote (left hand panels) and wildtype (right hand panels) Brutus offspring. Blue staining shows the localisation of *Basp1* expression. A,B, spleen and pancreas (x10); C,D, female reproductive tract (x25); E,F lungs (x10); G,H, thymus (x12.5); I,J, testis (x20); K,L, lymph nodes (K x80 and L x40); M,N, brain dorsal view (x8); O,P, brain ventral view (x8). s, spleen; p, pancreas; u, uterus; d, oviduct; ov, ovary; v, bronchial vessel; l, lymph node; f, fat; cb, cerebrum; t, thalamus; o, olfactory bulbs; cl, cerebellum; e, external capsule; m, medulla oblongata; on, optic nerve; sp, septum; p, pons.


p
s
A
s
p
B
u
d
ov
C
u
d
ov
D
v
E
F
G
H


I
J
K
l
f
l
L
M
cb
o
t
m
e
cl
N
O
sp
on
m
p
e
cl
P

density of structures within the cortex and medulla appeared to be very similar between both genotypes. Within the testes, it was clear that the seminiferous tubules of the heterozygous testes were immature, with no mature sperm present as in the wildtype testes. Also, the Leydig cells were small and sparse within the heterozygous testes leading to the appearance of a larger interstitial space.

Following the observation at dissection in some animals of air in the stomach and intestines, combined with the fact that none of the *Basp1*$^{+/-}$ mice survived past weaning age (~3 weeks), it was hypothesised that the cause of death might be due to a feeding problem arising following the withdrawal of milk. In order to investigate this possibility further it was decided to weigh one of these heterozygous mice daily, to observe its feeding behaviour and to check for overgrown teeth. During weaning all pups were given 'mash' food by animal house staff for one week before being taken away from their mother, or longer if the mice were struggling to thrive. Mash food is standard rodent food watered down to an almost liquid form. Observing feeding behaviours showed that these mice were attempting to eat this mash and therefore ingesting both water and nutrients. There was no apparent overgrowth of teeth and weight gain, although slow, did occur on a day by day basis. Following these observations it was also decided to check for cleft palate. Cleft palate is a condition in which the two plates that form the hard palate at the roof of the mouth are not completely joined, a condition that results in feeding difficulties (Halford et al., 2000). To check for cleft palate the Brutus chimera was set up for breeding with an F1 female and resulting day 1 pups were culled. A tail clip was taken from each pup in order that genotyping could be carried out, and the heads were fixed and embedded (see section 2.3.1). As the heads were large tissues for the fixative to penetrate, the skulls were first cut longitudinally down the centre and decalcified overnight (see section 2.3.6). Once fixed the heads were embedded nose down so that sectioning could be carried out to reveal cross sections through the palate. Serial sections up until the mid-point of the eye were taken and stained with H&E in order to visualise the structures (see section 2.3.4). Due to Brutus dying of old age, only one litter of pups was available to examine for cleft palate. Within the litter of six pups only one was heterozygous in genotype. Sections from this heterozygote and one wildtype were compared and an intact palate was

seen in both mice (Figure 4.11). This litter was also observed to check for milk in the stomach. All five wildtype pups had milk in their stomachs whereas it was not obvious that the heterozygous pup did. Total body weight and the weights of both the stomach and heart were taken from this litter (Table 4.7). The weight of the stomach was taken as a further indicator of feeding. The weight of the heart was taken to see if the differences in weight could be due to delayed growth, the hypothesis being that if *Basp1* is not expressed in the heart (as is the indication from the *lacZ* expression pattern) then the heart should also be proportionally smaller than the wildtype siblings if the defect is due to delayed growth and the same weight as wildtype siblings if not. Referring to Table 4.7 it is apparent that the total weight of the heterozygote is less that that of the wildtype siblings and the weight of the stomach was markedly less. The weight of the heart however was comparable to the wildtype siblings, and indeed it was not the smallest heart of the litter. Without further litters to examine it is not possible to draw strong conclusions from this weight data, so far however the data does not point to a delayed growth phenotype.

A final thought was the possibility that the mice were in fact homozygous rather than heterozygous knockouts, thereby providing a possible explanation for the neonatal death. Although the Southern hybridization results of the ES cell line that was used to generate Brutus chimera (Basp3, from 1st Transfection, Southern hybridization result not shown) showed that a targeted allele and a wildtype allele were present within this cell line, it was still possible that this line had been derived from a mixed clone; although unlikely it was therefore possible that this cell line contained a mix of cells some with one and some with two targeted alleles. To address this possibility PCRs were carried out on ear clip DNA from the dead and culled mice to amplify a product in exon 1 of *Basp1* [3F/3R] and exon 2 of *Basp1* [2F/2R]. Amplification of an exon 1 product should be possible with both a wildtype and targeted *Basp1* allele, whereas an exon 2 product should only be possible with a wildtype *Basp1* allele due to the deletion of this exon in the targeted allele. If two targeted alleles were present, as in the case of a homozygous deletion, no exon 2 product could be amplified. However, exon 1 and exon 2 PCR products were amplified from all dead and culled mice (results

Figure 4.11. Coronal sections of the heads of day 1 pups stained with haematoxylin and eosin in order to visualize the palate structures. (A) Wildtype littermate (x40); (B) Basp1+/- heterozygous knockout (x40). n, nose; t, tongue; p, palate.

A

B

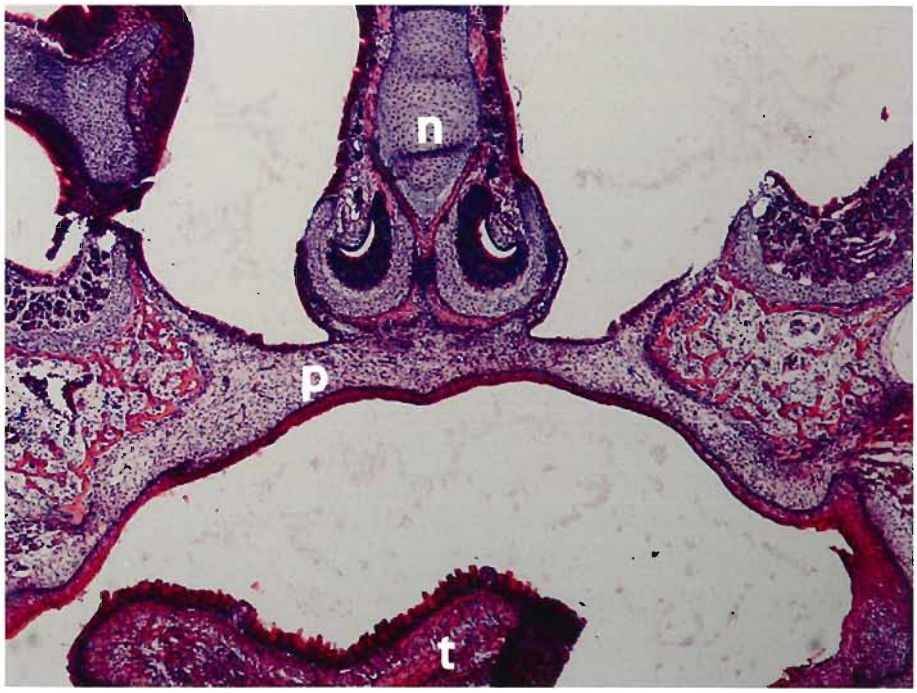


Table 4.7

pup number	sex	genotype	total weight (g)	stomach weight (g)	stomach relative weight (%)	heart weight (g)	heart relative weight (%)
1	♀	wt	1.5026	0.0704	4.69	0.0071	0.47
2	♀	wt	1.6874	0.1358	8.05	0.0110	0.65
3	♂	+/-	1.3947	0.0092	0.66	0.0077	0.55
4	♀	wt	1.5585	0.1201	7.71	0.0087	0.56
5	♂	wt	1.6258	0.0944	5.81	0.0087	0.54
6	♀	wt	1.5685	0.0804	5.13	0.0075	0.48

Table 4.7. Weight data of a single litter of day 1 *Basp1*$^{+/-}$ pups. Shows sex, genotype, total weight, stomach weight, heart weight and stomach and heart relative weights (weight as a percentage of total body weight) of a litter of six *Basp1*$^{+/-}$ pups dissected on the day of birth.

not shown). So in conclusion, the mice were heterozygous and not homozygous knockouts as there was one wildtype allele.

4.4.11 Geneticin-resistant ES cell clones with incorrectly targeted *Basp1* alleles

During the examination of geneticin resistant ES cell clones electroporated with the *Basp1* targeting construct it seemed that several of the clones were incorrectly targeted within the *Basp1* locus, as they gave bands of odd sizes on Southern blots probed with the *Basp1* specific probes C4 and C1.6 (Figure 4.4). It was a possibility that the integration of the *Basp1* targeted allele within the incorrectly targeted clones had resulted in a mutation resulting in a hypomorphic allele. A hypomorphic allele is one which can result in a reduction of gene expression rather than an absence of gene expression and protein product as would be the case with a null mutation (Rucker et al., 2000). Alternatively, incorrect targeting could have resulted in a mutation within the required gene without the genomic deletion of important regulatory regions. These clones could thereby provide a potentially less severe phenotype than a full null and were consequently investigated further in order to find out the nature of the molecular lesion at the *Basp1* locus.

RT-PCR analysis was carried out on five of these clones using primers in *Basp1* exon 1 and the SAβgeo cassette [3Fshort/βgalRL] designed to amplify a band only if the selection cassette of the targeting construct had been inserted downstream of *Basp1* exon 1 prior to exon 2 (results not shown). The parental cell line E14C4 and the correctly targeted cell line clone Basp65 were used as controls, and as expected a PCR product was derived from the targeted cell line (clone Basp65) and not from the parental cell line (E14C4). The RT-PCR analysis of clones Basp33, Basp39, Basp62, & Basp78 showed no amplification product suggesting that the construct was not inserted immediately downstream of exon 1 in these clones and therefore had not created a fusion splice product of exon 1 and the SAβgeo cassette, thereby confirming the Southern blot result that they were targeted in an unexpected manner. Clone Basp68 however did show an amplification product of the correctly targeted size. It was therefore assumed that

clone Basp68 was either correctly targeted and the results of the Southern blot had been misinterpreted due to background hybridization of the 3' probe (Figure 4.4D) or that the *Basp1* targeted allele had been inserted in two positions within the *Basp1* genetic locus. Clone Basp68 was therefore not a candidate for a hypomorphic allele.

The Southern blot was probed using the SAβgeo probe to verify that this cassette had indeed been inserted (Figure 4.4). Clones Basp62 & Basp78 contained one insertion of this cassette and clone Basp39 contained two. The odd clones were then tested for *lacZ* expression by exposure to X-gal in staining solution (see section 2.2.8). As the *lacZ* cassette provides an important diagnostic marker for use in knockout mice it was preferable that this cassette had remained intact in any clones that were to be used in the generation of these animals. It was found that clones Basp39, Basp62 & Basp78 did express β-galactosidase, whereas clone Basp33 did not. Although clone Basp33 lacked *lacZ* expression and did not contain an intact SAβgeo cassette as shown by Southern hybridization, it must have contained the neomycin resistance gene of this cassette in order for it to have survived G418 selection.

The two odd clones (Basp62 and Basp78) that contain one copy of the SAβgeo cassette and express *lacZ* could be useful if the expression of *Basp1* is disrupted in a way that unlike the expected targeting leads to a less severe phenotype. In an attempt to find the insertional location in order to predict what affect it might have *in vivo*, a further RT-PCR was carried out with primers designed to amplify a product between different regions of *Basp1* and the SAβgeo cassette. Primers used were 2FN (downstream of protein coding region in exon 2)/βgalRL (within the SAβgeo cassette) or 14F (3' end of exon 2)/βgalRL. Neither of these RT-PCRs gave a band, suggesting that the SAβgeo cassette had not been inserted downstream of these *Basp1* sites (at least not at the resolution distance for RT-PCR). It has to have been inserted somewhere downstream of the promoter however in order for the promoter-less βgeo to be expressed. Therefore it is still possible that the level of Basp1 protein expression could be affected in these clones leading to a hypomorphic allele, however it would take further

investigation to find either the exact site of DNA disruption, or an *in vitro* protein expression assay to check for *Basp1* expression levels.

4.5 Discussion

4.5.1 Efficiency of molecular cloning procedures

With all multistep molecular cloning strategies much prior thought needs to go into construct design to allow for the most efficient number of cloning steps. As is evident from section 4.4.1 detailing the construction of the targeting vector, not all of the cloning steps were straight forward.

Importantly, it was found that once the plasmids had reached a certain size (somewhere between 7.5-11.8kb) the efficiency of replication within the host bacterium fell dramatically. This drop in efficiency of replication was found to be the reason that step 5 of the cloning strategy was so problematic. Only the small parental plasmids lacking the SAβgeo cassette insert were being replicated. To overcome this problem, XL10-Gold® ultracompetent cells were used instead of the DH5α electrocompetent cells made within the laboratory (see section 2.1.6). The efficiency of these commercially made chemically competent cells being $5x10^9$ transformants / μg of pUC18 DNA compared to $5x10^7$ transformants / μg of pUC18 DNA of the laboratory made electrocompetent cells. Additionally, the XL10-Gold® ultracompetent cells allow more efficient transformation of large DNA molecules and exhibit faster growth than DH5α electrocompetent cells.

4.5.2 Efficiency of gene targeting procedures

Following transfection of the targeting construct into ES cells it was found that 30 out of 36 geneticin resistant clones (83%) were correctly targeted. This is a very high percentage out of such a small number of geneticin resistant clones tested. In comparison, when the *Wt1* gene was targeted by Kreidberg et al. (1993), only 29 out of 288 geneticin resistant clones (10%) were correctly targeted. Our high efficiency of targeting is most probably attributed to the three levels of selection: the positive selection of the neomycin resistance gene, the

negative selection of the diphtheria toxin gene and the splice acceptor allowing expression of the neomycin resistance gene only following integration downstream of an endogenous promoter. Other factors which most likely contributed to this high level of correct targeting were the length of the homologous arms and the isogenic nature of the DNA of the targeting vector and the host ES cell genome. The length of homology within a targeting construct is known to dramatically alter the targeting efficiency. It has been demonstrated with the targeting of the *hypoxanthine phosphoribosyltransferase* gene, that when the combined length of the homologous arms were increased from 1.3kb to 6.8kb that there was at least a 200-fold overall increase in targeting efficiency (Hasty et al., 1991). A step wise increase in efficiency was observed with lengths between these ranges, the most critical length being between 1.7kb and 1.9kb where additions of 225bp gave a >21-fold per kb increase in targeting efficiency. Lengths of homology below 1.7kb were insufficient for targeting events and above 4.2kb only gave a modest increase in efficiency of at least 1.6-fold. Additionally, it would seem that the total length of homology is a more important factor than the length of the shorter arm of homology; in their study there was no difference in efficiency when the shorter arm was between 472bp and 1.2kb. In the case of the *Basp1* targeting construct the 5' arm was 4.5kb and the 3' arm was 3.4kb, giving an overall length of homology of 7.9kb. This size should therefore give a high efficiency of correct targeting, indeed the arms could have been smaller but the position of appropriate restriction sites and therefore the ease of the cloning steps predetermined these sizes. As for the use of isogenic DNA, this is known to be important due to base pair polymorphisms between genomes of different mouse strains. One study comparing the efficiency of gene targeting in ES cells using homologous recombination in ES cells with isogenic (129) and non-isogenic (BALB/c) DNA of the *retinoblastoma susceptibility gene* showed a 20-fold increase with the isogenic DNA (Riele et al., 1992). In the case of the *Basp1* targeting the ES cells were derived from a 129 strain and the DNA used was of the same origin.

As outlined in the results section 4.4.4, once correctly targeted ES cell clones had satisfied several tests for suitability they were used for blastocyst injections. Not all chimeric mice generated gave germline transmission however; this isn't

unusual as not every host blastocyst will have ES cells that have contributed to the germline. Kreidberg et al. (1993) whilst targeting the *Wt1* gene documented that only one out of six ES cell lines transmitted the mutation through the germline. It was possible that the low percentage of euploidy exhibited in the two cell lines karyotypically tested (Basp55 and Basp65) were the reason these cell lines did not lead to germline transmission. Another possible reason could have been the loss of the Y chromosome which could not be specifically detected with the karyotypic method used; without the Y chromosome cell lines are less likely to transmit via the germline. Y chromosome loss has been reported in 2% of subclones following transfection procedures (Eggan et al., 2002). An additional procedure called Giemsa banding analysis could have been carried out in order to detect the presence of the Y chromosome in cell lines before using them in injection sessions, however, it was decided that because these indicators are not absolutely diagnostic, but act only as a guide as to which cell lines are the most likely to germline transmit, the decision was taken to inject several other 'promising' cell lines without embarking on further tests for suitability.

Additionally, higher numbers of injected blastocysts did not always lead to higher numbers of live offspring and higher numbers of live offspring did not always result in a high percentage of chimeras. One study has reported that the percentage of pups born following blastocyst injection and the percentage of resulting chimeras from two euploid ES cell lines (determined with chromosome counts and Giemsa banding analysis) that transmitted through the germline was between 68-70% and 62-65% respectively (Longo et al., 1997). However, this study also showed that one ES cell line with 100% euploidy had a low percentage birth rate (54%) and low appearance of chimeras (10%) and did not transmit through the germline, and also that three ES cell lines with only 50-60% euploidy had a low percentage birth rate (47-60%) and a low appearance of chimeras (40-45%) but did transmit through the germline. Only ES cell lines with a very low percentage of euploidy (<50%) never transmitted through the germline (Longo et al., 1997). There are therefore many possible reasons for our observations of low numbers of live offspring and of a low percentage of chimeras for some of the targeted ES cell lines, other than an abnormal karyotype. The first could be due to operator experience and practice of all the

techniques involved. This would undoubtedly have had an impact on the efficiency, however, since most of the cell lines were injected at various dates during the project, it is not possible to analyse the impact on a cell line by cell line basis. Secondly, even accurate implantation of injected blastocysts does not always result in birth if the pseudopregnant female mouse is not at the correct oestrus stage. There are ways to assess the receptiveness to implantation of the host mother (namely a vaginal plug check the morning after exposure to a male, and visualisation of a swollen ampulla and bloody ovaries) (Nagy et al., 2003), and these observations were adhered to when a female with these specifications were available. Mice with vaginal plugs were always available, whereas a swollen ampulla and bloody ovaries were not always seen. When females did not seem optimum for surgical implantation, a last resort would be to incubate injected blastocyst overnight in culture medium and implant into the oviduct of a 0.5day pseudopregnant female (as opposed to the uterus of a 2.5day pseudopregnant female). This technique was technically more difficult and of course there was no guarantee that optimum pseudopregnent females would be available the following day. A third possible reason for the difference in efficiency could have been the cell lines themselves. Although morphology (and even karyotype for certain lines) did appear to be optimum, if any chromosomal malformations existed, this could prevent the ES cells either from contributing to the embryos (resulting in no chimers, or at least not to the germline), or resulting in lethality of the chimera *in utero*. As discussed earlier, this is why a strategy of alternating between cell lines was employed. The most striking observation when considering efficiency was the low percentage of founder lines (only one out of 17 chimeric animals transmitted the *Basp1* KO allele through the germline). This fact leads me to conclude that the cell lines were to blame for the low efficiency, as other mouse knock out lines generated by myself and Dr Kim Moorwood (University of Bath) had much higher rates of transmission (see Chapter 5, Table 5.4).

The passage number of the parental cell line was between P18-P19 and of the targeted cell lines that generated chimeric mice was between P25-P35. These are high passage numbers, but not prohibitively so. A study carried out to investigate the effect of passage number on germline transmission in several ES cell line

strains (although not E14C4) has showed that with parental cell lines the maximum passage number is P26 and with targeted cell lines is P32 (Fedorov et al., 1997). Young parental ES cell lines of passage number <P15 were optimum for germline transmission. It is interesting that the passage numbers of the targeted ES cell lines may be higher than those of parental ES cell lines and still lead to germline transmission. This is due to sub-cloning of targeted ES cells. When a population of ES cells are maintained in culture, gradually certain clones loose their pluripotent potential and appear differentiated in culture; when these ES cell lines are transfected and sub-cloned, as long as an undifferentiated cell clone is selected, pluripotency of the line will be established to a higher level once again. The problem of the cell lines not transmitting through the germline in our circumstance must have occurred during the culturing and transfection stages; this is because the same parental cell line (E14C4) at the same passage number has been used by myself and Dr Kim Moorwood (University of Bath) to generate other germline transmitting mice.

4.5.3 Problems generating *Basp1* knockout founder lines

Possible reasons to explain the difficultly in generating founder lines of the targeted knockout were discussed in section 4.4.9 and attempts were made to clarify this issue. These reasons were: the heterozygous phenotype was lethal *in utero*; a second mutation linked to the first was present in the genome and causing neonatal death; a strain specific lethality for the mutation was being caused by modifier genes; or that the ES cells used to generate the chimera were of mixed origin. Whilst there is still a slim possibility that the ES cells were of mixed origin or that a second mutation linked to the first was preventing the birth of *Basp1*$^{+/-}$ mice, this by itself could only explain a reduced birth rate of the *Basp1*$^{+/-}$ mice, not their subsequent postnatal death. If a second mutation was present then it would be expected that a number of germline transmission wildtype mice would also die postnatally due to the inheritance of the second lethal mutation, but this did not happen. As pups of the heterozygous knockout genotype were eventually born this refutes that the phenotype was lethal *in utero*; it is however still possible that each mouse exhibited an abnormal phenotype of variable severity caused by the mutation, resulting in a number of *Basp1*$^{+/-}$ mice

dying *in utero*, decreasing the number of pups of this genotype observed by screening after birth. Any difference in phenotypic severity could be due to the presence of modifier genes; the alleles of nonlinked genes being inherited randomly.

4.5.4 Neonatal death of *Basp1*$^{+/-}$ mice

It was interesting that of the *Basp1*$^{+/-}$ pups that survived past birth all either died or had to be culled due to ill health around weaning age (~3 weeks after birth). Due to this observation and the presence of air in the stomach and guts of these ill mice, it was hypothesised that these mice die from problems either of eating, drinking or digesting food. From the length of time these mice survived it was apparent that as pups they were able to feed on milk as they would have died shortly following birth if this was not the case. It was also observed that these mice during weaning ate 'mash' food. The next hypothesis was that although these mice were attempting to feed they were having problems in doing so making it difficult to take in enough nutrients. To address these issues, mice were examined for overgrowth of teeth and day 1 pups were culled in order to check for cleft palate. Neither overgrowth of teeth or cleft palate was apparent. Unfortunately we were only able to examine one *Basp1*$^{+/-}$ pup for cleft palate, however the pup did not show signs of this malformation; of course it is impossible to know whether this pup would have survived past weaning age, but this would have gone against the precedent that had been observed up to that point. These positive feeding observations combined with *in situ* hybridization (see section 3.4.7) and immunohistochemistry results (Carpenter et al., 2004) showing high levels of Basp1 expression in the gut leads us to conclude that the problem is more likely to be with digestion of food and nutrient uptake due to a physiological defect of the gut in the absence of Basp1. Alternatively, feeding could still be the reason for postnatal death possibly due to a malformation of the tongue, another organ showing high levels of *Basp1* expression with immunohistochemistry results (Carpenter et al., 2004), or due to behavioural reasons caused by a malformation within the brain. Without the establishment of a founder line, thereby providing the means to examine further mice, it was not

possible to determine conclusively the reason for the postnatal death of these *Basp1* deficient mice.

4.5.5 Comparisons with a previously described *Basp1* knockout mouse line

As previously mentioned, the existence of a *Basp1* KO mouse published by Frey et al. (2000) was unknown at the commencement of the project and therefore the viability of this knockout strain was impossible to accurately predict; as it happens their publication shows that the heterozygous knockout mice, and a proportion of the homozygous null mice were viable. The severity of the phenotype observed in the heterozygous *Basp1* mice in this study was therefore surprising in light of the published mouse line being viable (Frey et al., 2000). In the paper by Frey et al. (2000) $Basp1^{+/-}$ and $Basp1^{-/-}$ mice were born with the expected Mendelian ratios. Approximately 66% of $Basp1^{-/-}$ mice died between birth and 4 months but only a small number of $Basp1^{+/-}$ mice died postnatally. The surviving $Basp1^{-/-}$ adults were 50-70% the weight of their wildtype litter-mates and the surviving $Basp1^{+/-}$ mice were 80% of normal weight, but no other phenotypic characterisation was mentioned. Since the heterozygous knockout mice do not have a wildtype phenotype and it has been shown that these mice express approximately a third of wildtype Basp1 levels, Basp1 was said to be haploinsufficent (Frey et al., 2000).

The severe phenotype observed in our *Basp1* heterozygous knockout mice in comparison to the previously published *Basp1* KO phenotype could be due to several factors. The most obvious would be that any strain differences (note that strain information was not provide by Frey et al. (2000)) could lead to the differences seen, as this is a well documented phenomenon (Doetschman, 1999). Indeed, the phenotype of the *Wt1* KO mouse model has been shown to be affected by strain differences. On the C57BL/6:129/Sv hybrid background heterozygous knockouts did not display an overt phenotype (Kreidberg et al., 1993), whereas on a mixed genetic background (CBA, SWR, 129/Ola, C57BL/6, BalB/c) 40% of heterozygous knockouts died by 400 days after birth (Kreidberg et al., 1999) and on an inbred 129/Sv background female heterozygous knockouts were infertile (Kreidberg et al., 1999). The ES cells used by Frey et al.

(2000) are almost certainly to have been derived from 129 strain as in our case (this is because although C57BL/6 ES cells were derived in 1991 (Ledermann and Burki, 1991) they perform their role poorly with less germline transmission seen compared to 129 ES cells (Auerbach et al., 2000)); the host blastocyst strain and the strain of mouse used to breed with the resulting chimeras could however have been different. Another possibility is that whilst our mutation was designed to eliminate all Basp1 protein, the knockout published by Frey et al. (2000) could have been designed in another way that did not eliminate the full length protein thereby reducing the severity of the phenotype (note that this information was not provide by Frey et al. (2000)). It would be possible to eliminate some of these possibilities by generating further chimeras and analysing the founder offspring to compare two or more independently generated lines. This would involve a lot more time and resources, especially as it would be advisable to go back to the stage of transfecting ES cells and reselecting correctly targeted clones before embarking on further rounds of blastocyst injections.

4.5.6 Comparisons with the *Gap43* knockout mouse line

A striking comparison can be drawn between our *Basp1*$^{+/-}$ mice and the *Gap43*$^{-/-}$ KO mice reported by Strittmatter et al. (1995). 90% of these *Gap43*$^{-/-}$ null mice died between birth and weaning (45% between PD 0-2, 45% around weaning age). Also, weight gain was poor with average body weights being 55% that of wildtype littermates. Necropsy of these mice showed the stomach and gut were devoid of food as with our *Basp1*$^{+/-}$ mice that were examined to check for this possibility - this absence of food therefore suggesting starvation. Organs examined during necropsy of the *Gap43*$^{-/-}$ mice did not show any abnormalities, just a proportional decrease in body and organ weights which was also the case with our *Basp1*$^{+/-}$ mice. These two proteins, although not sharing a high sequence homology (Maekawa et al., 1993; Wiederkehr et al., 1997) are functionally similar; indeed Gap43 can substitute for Basp1 *in vivo* as shown by the normal phenotype of *Basp1*$^{gap43/gap43}$ mice (see section 1.5.6) (Frey et al., 2000). This suggests the possibility of a similar phenotype for these two knockout mouse strains.

4.5.7 Sites of *Basp1* expression

The *lacZ* reporter within the targeting vector allowed for the sites of *Basp1* expression to be traced; this is because any sites of β-galactosidase expression upon X-gal staining should correspond to the sites of *Basp1* expression due to the *lacZ* gene being driven by the endogenous *Basp1* promoter. β-galactosidase expression was monitored in *Basp1*$^{+/-}$ mice at E15.5 and neonatally at ~3 weeks of age. Some of these sites of *Basp1* expression correspond to the already documented sites of embryonic protein expression (Carpenter et al., 2004) and to the sites of embryonic expression observed with *in situ* hybridization (see section 3.4.7). Previously undocumented sites of expression include the pancreas, lymph nodes, bronchial vessels of the lung, vertebrae and umbilical cord. The staining seen in the pancreas and lymph nodes were from dissected organs of the neonate (Figure 4.10) and are believable sites of expression as there was no staining seen in these organs of a wildtype litter mate. The staining seen in the vertebrae and umbilical cord are less reliable as they were visualized as whole mount staining in the E15.5 embryo (Figure 4.7) and therefore individual organs were harder to identify; it is also possible, particularly with the umbilical cord, that the staining could have been only in the epithelial linings. Interestingly there was no expression seen in the heart of either the neonate or the E15.5 embryo, even though previous reports have identified expression in this organ by immunohistochemistry (Carpenter et al., 2004) The *lacZ* staining pattern corresponds well with the *in situ* hybridization pattern at E15.5 (see section 3.4.7) which also did not show *Basp1* RNA expression in the heart.

4.5.8 Possibilities for other *Basp1* transgenic mouse models

As it remains unresolved whether the phenotype of the *Basp1*$^{+/-}$ mice is a true phenotype caused by the loss of Basp1 protein on the 129:C57BL/6 hybrid strain background or caused by another additional mutation, and the only germline transmitting chimera is dead, there are several further future possibilities to resolve this:

- To create further chimeras either from the other correctly targeted ES cell lines that have been generated or from further newly generated ES cell lines, possibly using different blastocyst host strains or breeding strains

- To request the previously published *Basp1* KO mice (Frey et al., 2000) be sent to us so that we could carry out work involved in kidney development and enter into a collaboration
- To design a conditional knockout
- To investigate further the possibility that the 'odd size' ES cell clones are suitable to generate a mouse model with a *Basp1* hypomorphic allele

The first possibility as long as the chimeras were to be generated from already correctly targeted ES cell lines, could begin within three months once the mouse colony from which the host blastocysts were to be derived are bred to reach adequate stock numbers. If further ES cell lines were to be generated, now that the experimental procedures have been optimised, this could take approximately 3 months presuming correctly targeted clones were found within the first round of transfection.

The second possibility seems the most labour saving although we do not know if the published *Basp1* KO mice still exist, or if collaboration could be entered into. We have also been unable to ascertain exactly which region of the *Basp1* gene had been deleted or what strain background these mice were maintained on. Additionally, often when mice are needed by our animal facility they need to be re-derived from embryos as the status of our facility is free of many pathogens that could be introduced into the colonies from outside sources. So this in itself may take time and several attempts to re-derive the stock.

The third possibility, to generate a conditional knockout animal, would mean revised cloning steps before further targeting of ES cells and the subsequent blastocyst injection of embryos could begin. As already discussed above, multistep cloning strategies need much thought and practically can be a lengthy procedure. The advantages would be that the phenotype could be observed in only one organ i.e. the kidney, and that any lethal affects that arise due to the absence of this gene in other organs could be circumvented to allow the organ of interest to be studied at a later stage in the life of the animal. As only one conventionally targeted ES cell has so far been examined it may be too early to embark on this conditional strategy.

The final possibility presuming that one of the incorrectly targeted ES cells exhibited a hypomorphic allele would be a cheaper and less labour intensive alternative to designing a conditional knockout. Again, this would only be appropriate assuming conventional targeting does cause lethality in heterozygous knockout mice and a less severe phenotype was needed. Further investigation into the 'odd size' clones would need to be carried out to find either the exact site of DNA disruption or to check for reduced protein levels.

4.6 Appendix

C9 N°.2 glycerol stock made on 16/10/03, maxiprep made on 17/11/03.

Generation of a mouse model to study the role of Wt1-AS

5.1 Summary

This chapter will outline the strategy used to generate two mouse models to investigate the role of the *Wt1-AS* gene *in vivo*. One of the two mouse models was successfully generated within the time frame of this thesis. The phenotypic analysis of this model will be presented, specifically the histopathological affect this genetic alteration had on organ development and maintenance.

Within this chapter the generation of the targeting construct via molecular cloning, expansion of geneticin resistant clones following electroporation, identification of correctly targeted clones via Southern hybridization/PCR and the generation of transgenic mice via blastocyst injection was carried out jointly by the author (Joanne Allsop) and Dr Kim Moorwood (University of Bath). Preliminary cloning steps carried out by other persons will be acknowledged within the text. Other experimental procedures within this chapter, including all the phenotypic characterisation, were carried out solely by the author (Joanne Allsop).

5.2 Introduction

WT1-AS is an overlapping antisense transcript of *WT1* that produces an RNA that is not translated into protein. Several pieces of evidence point to *WT1* and *WT1-AS* being able to reciprocally regulate each others expression by positive feedback. Firstly, antisense transcription is coordinate with sense mRNA and protein expression, and is strictly temporally and spatially regulated in both normal kidney (Moorwood et al., 1998) and Wilms' tumours (Yeger et al., 1992). Secondly, transfection studies have shown transactivation of the *WT1-AS* promoter by WT1 protein *in vitro*. Specifically, when cell lines expressing elevated levels of *WT1* sense mRNA and protein were transiently transfected with a reporter construct expressing the *WT1-AS* promoter, expression from the *WT1-AS* promoter was elevated (Malik et al., 1995). This transactivation was found to be WT1 isoform specific (Moorwood et al., 1999). Thirdly, *WT1-AS*

positively modulates WT1 protein levels *in vitro*. When cell lines expressing elevated levels of *WT1* mRNA and protein were stably transfected with a reporter construct expressing the *WT1-AS* exon 1 region, in the majority of cases WT1 protein levels were elevated (Malik et al., 1995; Moorwood et al., 1998).

In order to investigate this positive feedback between *WT1* and *WT1-AS in vivo*, it was deemed necessary to generate a mouse model lacking *WT1-AS*. As the *WT1-AS* gene is not translated into protein, there was no protein coding region to knock-out. Instead, one option would have been to delete the promoter region of this gene. The promoter region of the human *WT1-AS* transcript was found to be located in intron 1 of *WT1* (Malik et al., 1994), however, the promoter region of the murine *Wt1-AS* gene was unknown. Preliminary experiments to locate the murine *Wt1-AS* promoter were unsuccessful (see section 5.4.1). It was therefore decided to eliminate the *Wt1-AS* transcript by another means as described below.

Analysis of human and mouse sequence data has highlighted two upstream conserved regions, termed MCRI and MCRII which lie ~11-12kb and ~4-5kb upstream of *WT1* exon 1, respectively (see Chapter 1, Figure 1.3). Both elements are located at the 3' termini of alternately spliced, non-coding antisense transcripts from human kidney. MCRII was the most conserved sharing 94% sequence identity between mouse and man over 289 nucleotides. At least two *WT1-AS* spliceoforms are known to be transcribed through the MCRII region within the human genome (Dallosso et al., 2007). It is also known that the MCRII region is transcribed within the murine genome providing evidence of the potential functional importance of MCRII; RT-PCR was carried out using random primed cDNA made from foetal mouse kidney RNA using primer pairs designed within the MCRII region. Although using random primed cDNA is not strand specific, the MCRII region is located within a region of the genome that is not known to contain any overlapping sense transcripts, therefore, any product should be specific for *Wt1-AS* transcription (personal communication, CLIC Sargent Research Unit, University of Bristol).

The high inter-species homology and other circumstantial evidence pointing to a regulatory role for MCRII as a DNA element (see section 1.4.10) lead to this

region being chosen as the site for a mouse model genomic deletion. In addition, a second mouse model containing a knock-in of a transcriptional terminator sequence was also generated in order to investigate the phenotype caused by truncated expression of the *Wt1-AS* transcript. The process of terminating antisense gene expression by the positioning of a transcriptional terminator sequence within the genomic locus was successfully achieved by Sleutels et al. (2002) for the non-coding *Air* RNA (see section 1.3.1.3).

The goal of deleting and truncating the MCRII region was to dissect the role of the MCRII enhancer region in *Wt1* gene expression and to investigate whether *Wt1-AS* transcription of this region is essential for normal *Wt1* expression and hence mouse development. Although the generation of these two mouse models (the MCRII targeted deletion and the *Wt1-AS* truncation allele) does not permit a direct examination of the role of the full length *Wt1-AS* transcript, they do provide the means to answer the following questions:

- Does MCRII function at the level of RNA as part of *Wt1-AS* transcripts, at the level of DNA as a transcriptional enhancer, or as both?
- Does MCRII mediate regulation and distribution of *Wt1* sense RNA and/or Wt1 protein levels during mouse development?
- Does MCRII mediate alterations in *Wt1-AS* expression levels during mouse development?
- Could mutations at MCRII produce hypomorphic *Wt1* alleles that model developmental defects seen in Denys-Drash and Frasier syndrome, by disrupting MCRII?

We aim to answer these questions by examining the two mouse models for phenotypic differences in organ histology, and by monitoring expression of *Wt1*, *Wt1-AS*, and WT1 levels. The MCRII targeted deletion model will express a truncated version of the *Wt1-AS* transcript and the genomic MCRII region will be deleted. The *Wt1-AS* truncation allele model, will express a truncated version of the *Wt1-AS* transcript but the genomic MCRII region will remain intact. The models will allow the role MCRII plays in mouse development to be dissected into DNA enhancer or RNA-mediated activities.

The two required genetic alterations were first generated within ES cells. A parental targeting construct was introduced into ES cells using the process of homologous recombination. This targeting was followed by *cre recombinase* (*cre*)-mediated recombination between loxP sites within the targeted locus in order to generate separate ES cell clones containing the two distinct alleles from a single targeted integration. The targeted ES cells that had undergone correct recombination events were then used to generate mouse models by the process of blastocyst injection.

5.3 Methods

Many experimental procedures relevant to this chapter have already been described in Chapter 4. Specific details have been outlined below.

Plasmids generated prior to the commencement of molecular cloning within this chapter can be found in Table 5.1. The sequences of oligonucleotide primer pairs used to generate linkers during the molecular cloning steps can be found in Table 5.2. Primer sequences used for PCR or RT-PCR within this chapter can be found in Table 5.3.

5.3.1 Introduction of the targeting construct into ES cells

Transfection of the construct into embryonic stem cells was carried out using electroporation in the same way as the *Bsp1* targeting construct (see section 4.4.2); the same parental ES cell line with the same culturing and transfection conditions were used. Both the antibiotic dose response test and the test to ascertain the optimum electrical pulse conditions carried out in section 4.4.2 were essential for the successful transfection of the MCRII targeting construct. To reiterate, 225μg/ml of G418 was used in transfection experiments and electroporation was carried out with a single electrical pulse of 300V, 250μF.

25μg of *NotI* linearized targeting construct mcr2 gtv#2 was transfected with E14C4 ES cells (Smith, 1991) and grown in geneticin selective media (see section 2.2.6). The passage number of the parental E14C4 ES cell line was P19.

Table 5.1

resource/plasmid name	reference/source	description
Wt1 BAC	unpublished: kind gift of Nick Hastie, MRC, Edinburgh, UK	source is the 129 mouse strain
pMC1DTA	Yagi et al., 1990	pUC19 backbone, containing a DTA cassette driven by an MC1 promoter and terminated by SV40 DNA lacking a polyadenylation signal
LoxP-tPA	unpublished: kind gift of Marko Horb, McGill, Quebec, Canada	sub-cloned from the pGK neo tpA lox2 plasmid. containing the following adjacent sites: loxP site, a triple polyadenylation signal to prevent transcriptional read-through, loxP site.
pGK neo tpA lox2	Soriano, 1999	containing the following adjacent sites: slice acceptor site, loxP site, pGK promoter, neomycin resistance cassette, a triple polyadenylation signal to prevent transcriptional read-through, loxP site, *lacZ* gene, and polyadenylation sequence.
pCAG-CRE	Araki et al., 1995	pUC13 backbone, with a CMV enhancer/chicken β-actin (CAG) promoter driving *Cre* expression, terminated by a rabbit β-globin polyadenylation signal.

Table 5.1. Details of the DNA elements used within the MCRII knockout cloning strategy that were generated prior to commencement of the MCRII knockout project. Shows the reference or source of the WT1 BAC and each plasmid and a brief description of the components that make up each resource.

Table 5.2

oligonuleotide primer pairs	sequences
PACLOXF	5'ATAACTTCGTATAATGTATGCTATACGAAGTTATAT3'
PACLOXR	5'ATAACTTCGTATAGCATACATTATACGAAGTTAAT3'
HSLOXF	5'AGCTATAACTTCGTATAATGTATGCTATACGAAGTTATCCCGGG3'
HSLOXR	5'TCGACCCGGGATAACTTCGTATAGCATACATTATACGAAGTTAT3'
fab1	5'GGAGCTGGGGAACCCGTGGTGTGGCGTCGACACACTGCGGCTCCA GGCACCAGGAA 3'
fab2	5'TTCCTGGTGCCTGGAGCCGCAGTGTGTCGACGCCACACCACGGGT TCCCCAGCTCC3'

Table 5.2. Oligonucleotide primer pairs used to form the linkers during the MCRII knockout cloning strategy.

Table 5.3

primer pairs	sequences	Expected product size	annealing temp
WT1 Ex1	5'GAGGAGCAGTGCCTGAGCG3'	485bp	60°C
WT1 In1.1	5'TTCCGAAATCTCTAGGACACC3'		
WT1 In1.2	5'CTTGAGGAGAAACCTGGAGCG3'	252bp	60°C
WT1 In1.3	5'GTTTGGCCTATAACTTCTGGC3'		
WT1 In1.4	5'CTATCCCTTGGCAACTGAGT3'	371bp	60°C
WT1 In1.5	5'GGCTGATTAGGAACCCGCTG3'		
WT1 In1.6	5'CTTTCCTCACTTCAGATACAG3'	345bp	60°C
WT1 In1.7	5'GTTTCTTTGGAGACTCAGTTATG3'		
WT1 In1.8	5'GTCAGCAGAGGAAACCCAGTC3'	358bp	60°C
WT1 In1.9	5'AAGCAGGGGCACTAGCAGCG3'		
WT1 In2.1	5'GATGCGGGTCTTTACTGGATG3'	287bp	60°C
WT1 In2.2	5'GCGAGTCACTGGCAGAGGTTC3'		

WT1 In3.1 WT1 In3.2	5'CCTCCAAGGTTTACTATCTC3' 5'CAGTTAGAAACACAGCCAC3'	330bp	60°C
WT1 In4.1 WT1 In4.2	5'ATTAGTGCGGCGATGATTTG3' 5'GGAATCTGAGAAGTGGTTTG3'	358bp	60°C
WT1 Ex1 WT1 Ex2	5'GAGGAGCAGTGCCTGAGCG3' 5'TTTGAAGGAATGGTTGGGGAAC3'	280bp	60°C
RCmcr2DN mcr2 S	5'GATTTACGGGATAATAATAGAAG3' 5'CTGGTGCGAGATGCAAATTTTGC3'	variable	62°C
mcr2 G neolong	5'CCAGGTGTAGCATAAGTTCTC3' 5'ATCGCCTTCTATCGCCTTCTTG3'	~1kb	62°C
mcr2 HF mcr2 S	5'CCCTCTGAACCCCACTCCACC3' 5'CTGGTGCGAGATGCAAATTTTGC3'	variable	66°C
mcr2 CNLong mcr2 B	5'AGATGGTGCGCTCTGGACTGG3' 5'GACGCCACACCACGGGTTCC3'	243bp	60°C
mcr2 A mcr2 B	5'GCAAAGTTGAATCTGGGGTCAGG3' 5'GACGCCACACCACGGGTTCC3'	297bp	58°C
A11 A12	5'GACCCAGATCATGTTTGAGACC3' 5'GTTGGCATAGAGGTCTTTACGG3'	620bp (DNA) 530bp (RNA)	60°C
mcr2 SF mcr2 VR	5'GCAAAATTTGCATCTCGCACC3' 5'GCTCAGTTTCCAGACTAGCGC3'	334bp	60°C
mcr2 IF mcr2 JR	5'CTGACAGATCGCAAGGCTACAG3' 5'GAGAAAATATGACAGTAGCGTTC3'	490bp	56°C
endoL-FOR endoL-REV	5'GTCAAATGCGAGTGACAACGC3' 5'GCTCCCTACTGCTTCTGGTG3'	402bp	50°C
endoR-FOR endoR-REV	5'GCATCTTTGAGGACGGAGAG3' 5'TGAGACCAGGGAACCAACTC3'	238bp	55°C
βgeoR mcr2 S	5'CTCGTCCTGCAGTTCATTCA3' 5'CTGGTGCGAGATGCAAATTTTGC3'	~1.69kb	58°C

endo neolong	$^{5'}$CAGGGACTTTTCTAACTGGGG$^{3'}$ $^{5'}$ATCGCCTTCTATCGCCTTCTTG$^{3'}$	~1.98kb	58°C
mcr2 HF mcr2 GR	$^{5'}$CCCTCTGAACCCCACTCCACC$^{3'}$ $^{5'}$GAGAACTTATGCTACACCTGG$^{3'}$	961bp	50°C
mcr2 C ΔneoR	$^{5'}$AGTTCTAATTGTAACCTCAGA$^{3'}$ $^{5'}$CATCAAGCTGATCCGGAACC$^{3'}$	337bp	55°C
mcr2 TF mcr2 S	$^{5'}$CCCAGTTCTGCCTGGTCTTTC$^{3'}$ $^{5'}$CTGGTGCGAGATGCAAATTTTGC$^{3'}$	variable	58°C
WAS1-LONG ΔneoR	$^{5'}$TGTTCTGCAGTTACACTACTTATC$^{3'}$ $^{5'}$CATCAAGCTGATCCGGAACC$^{3'}$	variable	55°C

Table 5.3. Primer pairs used to perform PCR or RT-PCR. Shows primer names, sequences, the size product expected and the annealing temperature of the primer pairs.

5.3.2 Generation of DNA probes for Southern hybridization

The endoL and endoR probes were made by cloning PCR products from the amplification of the WT1 BAC with primers EndoL-FOR/EndoL-REV and EndoR-FOR/EndoR-REV.

The neo probe was made by digesting the SAβgeo plasmid (Friedrich and Soriano, 1991) (see Table 4.1) with *EcoRI* and *NcoI* and gel isolating the ~650bp fragment (Figure 5.4B).

5.3.3 Transfection of *Cre recombinase* into the correctly targeted ES cell clones

In order to create *Cre*-loxP deletions within the parental targeting construct, C100 P28-29 (P9-P10 following transfection) was transfected with a plasmid expressing *Cre recombinase* under the control of a modified chicken β-actin promoter (plasmid pCAG-CRE (Araki et al., 1995)). 20μg of circular plasmid was used. Electroporation conditions were the same as for the first round of transfection. As there was no antibiotic selection cassette within this plasmid, following transfection the cells were plated at a low density (50, 250, and 1000 cells per petri dish (35mm diameter) based on the pre-transfection density) and after a couple of days of recovery individual colonies were picked and transferred into separate wells of a 96-well plate.

5.4 Results

5.4.1 Searching for the *Wt1-AS* promoter region

In an attempt to find the 5' end of the *Wt1-AS* gene and hence narrow down the potential region of the *Wt1-AS* promoter, RT-PCR was carried out on E15.5 foetal kidney RNA with an array of primers designed in the intronic regions of the *Wt1* gene in order to ascertain how far upstream the *Wt1-AS* transcript extended. cDNA was made in two ways: using an antisense specific primer designed in exon 1 of *Wt1* [WT1 Ex1]; or by using oligodT primers (see section 2.1.19). There were advantages and disadvantages of both methods: using an antisense specific primer positively selects for cDNA of these low abundance

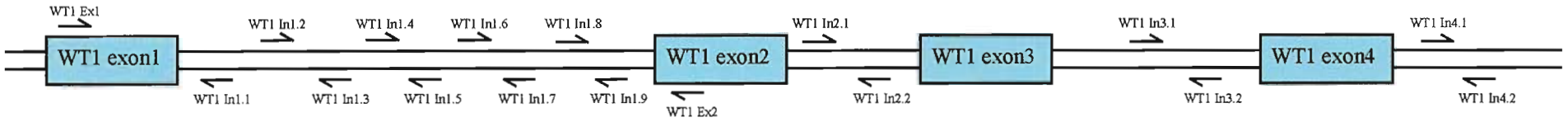
transcripts, however the primer chosen may not be very efficient; using oligodT primers was an efficient method of generating cDNA templates but it was not specific and the successful synthesis of cDNA corresponding to the *Wt1-AS* transcript would only occur if *Wt1-AS* has a polyadenylation signal (a fact which is unknown) or if it contains a string of A-base repeats. RT- cDNA made without the addition of reverse transcriptase to the first-strand synthesis reactions was used as a negative control for each reaction. RT-PCR products could only be amplified from the RT- cDNA if there was contaminating genomic DNA that was acting as a template for PCR.

Once the cDNA had been made using both the first-strand synthesis methods mentioned above, RT-PCR was carried out with primers designed to amplify the 5' region of *Wt1-AS* (Figure 5.1A and Table 5.3 for the primer sequences). RT-PCR products were made using both strand specific cDNA and oligodT cDNA with all but the first set of primers [WT1 Ex1/WT1 In1.1] where only the oligodT cDNA gave an RT-PCR product (Figure 5.1B). The RT-PCR results suggest that *Wt1-AS* either does have a polyadenylation signal or contains strings of A-base repeats. The absence of a product with the strand specific cDNA and the first set of primers can probably be put down to the fact that the cDNA was made with one of the same primers used in the RT-PCR reaction, this sequence therefore not being synthesised in its entirety. The absence of RT-PCR products in the RT- lanes show that the products were indeed amplified from cDNA and not genomic DNA.

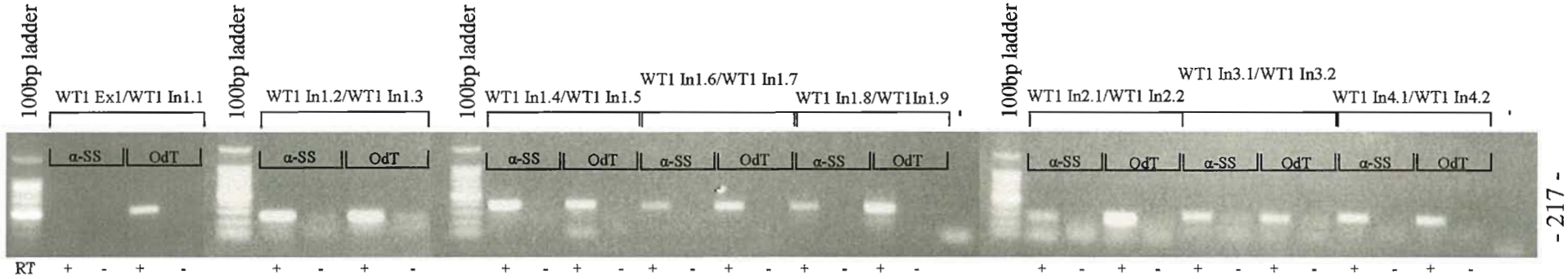
Mapping the mouse *Wt1-AS* transcript by strand specific RT-PCR within several introns of *Wt1* suggested that the 5' end of the transcript was not within intron 1 of *Wt1* (as in the human locus) but further downstream. Although the RT-PCRs seemed to be working well, the first indication of a problem occurred when products were still being amplified with primers designed in intron 3 [WT1 In3.1/WT1 In3.2] and intron 4 [WT1 In4.1/WT1 In4.2] of *Wt1* using cDNA made with the strand specific primer WT1 Ex1. The cDNA synthesis kit used (SuperScript™ II MMLV RT) should only be able to amplify cDNA products of between 0.5kb-10kb and in order to have a RT-PCR product with intron 4 primers, the cDNA synthesis reaction would have had to amplify over 15kb, a

Figure 5.1. RT-PCR to map the 5' end of the *Wt1-AS* transcript. (A) Schematic of the 5' Wt1 locus showing WT1 exons 1-4 and the relative position of the primer pairs designed around this locus. (B-D) Shows RT-PCR results carried out with a range of primer pairs on cDNA made from total RNA using either an antisense specific primer (α-SS), oligodT primers (OdT), a sense specific primer (SS) or no primer (none) with the SuperScript™ II Reverse Transcriptase Kit. The primer pairs used for each RT-PCR reaction are labelled above each well. RT-PCR reactions carried out with cDNA made from first strand synthesis reactions with and without reverse transcriptase are labelled + and – respectively.

A

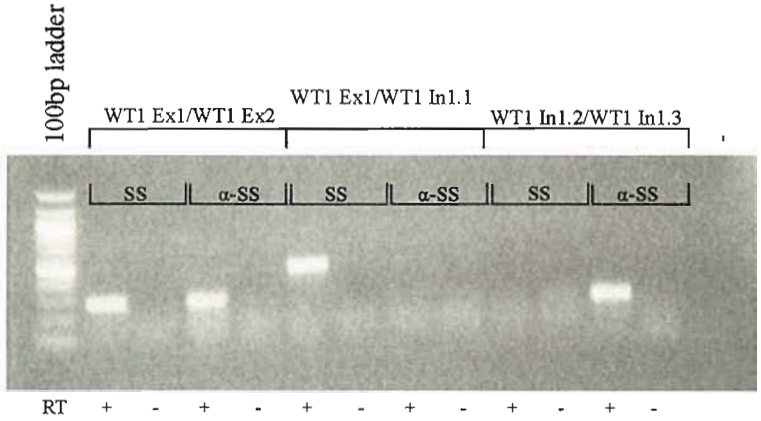

WT1 Ex1
WT1 In1.2
WT1 In1.4
WT1 In1.6
WT1 In1.8
WT1 In2.1
WT1 In3.1
WT1 In4.1
WT1 exon1
WT1 exon2
WT1 exon3
WT1 exon4
WT1 In1.1
WT1 In1.3
WT1 In1.5
WT1 In1.7
WT1 In1.9
WT1 Ex2
WT1 In2.2
WT1 In3.2
WT1 In4.2

B

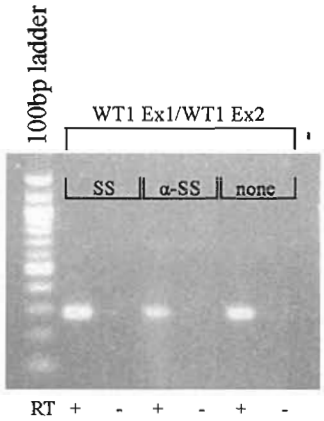

100bp ladder
WT1 Ex1/WT1 In1.1
α-SS
OdT
100bp ladder
WT1 In1.2/WT1 In1.3
100bp ladder
WT1 In1.4/WT1 In1.5
WT1 In1.6/WT1 In1.7
WT1 In1.8/WT1In1.9
100bp ladder
WT1 In2.1/WT1 In2.2
WT1 In3.1/WT1 In3.2
WT1 In4.1/WT1 In4.2
RT + - + - + - + - + - + - + - + - + - + - + - + - + -

- 217 -

C


100bp ladder
WT1 Ex1/WT1 Ex2
WT1 Ex1/WT1 In1.1
WT1 In1.2/WT1 In1.3
SS
α-SS
SS
α-SS
SS
α-SS
RT + - + - + - + - + - + -

D


100bp ladder
WT1 Ex1/WT1 Ex2
SS
α-SS
none
RT + - + - + -

distance much further than expected. To make sure that the RT-PCR products amplified were indeed antisense specific and not products amplified from cDNA made from *Wt1* pre-mRNA, sense specific cDNA was made from E15.5 foetal kidney RNA using a sense specific primer designed in intron 3 of *Wt1* [WT1 In3.2] and further RT-PCR reactions were carried out (Figure 5.1C). Three sets of primers were used: primers designed in exon 1 and exon 2 of *Wt1* [WT1 Ex1/WT1 Ex2]; primers designed in exon 1 and intron 1 of *Wt1* [WT1 Ex1/WT1 In1.1]; primers designed in intron 1 of *Wt1* [WT1 In1.2/WT1 In1.3]. The latter two primer pairs should give a product with both sense and antisense specific cDNA of 485bp and 252bp respectively if the sense specific cDNA was indeed amplifying *Wt1* pre-mRNA, presuming it was possible to synthesize a large enough sense specific cDNA from intron 3 to intron 1 (a size also over 15kb). The absence of the 485bp product with the antisense specific cDNA can be put down to the same circumstance as mentioned earlier, that the cDNA was made with one of the same primers used in the RT-PCR reaction and this sequence was therefore not synthesised in its entirety. The absence of the 252bp product with the sense specific cDNA cannot be accounted for in the same way, and has to be put down to either the cDNA synthesis product not spanning the intron1 region, or that sense specific cDNA was not made from the pre-mRNA. However, what is interesting is the presence of a 277bp product with both the sense and antisense specific cDNA using the exon1/exon2 primers. With both cDNA templates the product size should have been too large to amplify as the primers span intron 1 of *Wt1*, a total distance of ~5.8kb. The 277bp product actually amplified is the size expected for spliced sense cDNA that has no intronic sequence to span. Obviously an unexpected problem was occurring and to address this, the cDNA was resynthesized along with a control first-strand synthesis reaction with no primer. RT-PCR reactions with the exon1/exon2 primers were repeated (Figure 5.1D). The 277bp product was again amplified with the sense and antisense specific cDNA, but shockingly also with the cDNA made with no primer template; consequently, none of the RT-PCR results can be trusted to have been antisense specific, therefore the 5' end of the *Wt1-AS* transcript was no closer to being found.

A literature search for the phenomenon of cDNA being synthesised without the addition of a primer to the first-strand synthesis reaction shows that this is not the first documented case of the problem occurring (Agranovsky, 1992). Indeed, Ambion® has produced a kit called EndoFree RT™ to combat the problem of reverse transcription products being generated without the addition of primers by a process called 'endogenous priming'. The kit works by using a reverse transcriptase that reduces endogenously primed cDNA products. This kit was used with poly(A)$^+$ RNA extracted from E17.5 mouse kidney (see section 2.4.2) to try to combat endogenous priming around the *Wt1-AS* locus. Oligo(dA)$_{80}$ primer was also added during the first strand synthesis stage in order to trap any olido(dT) primer that may not have been removed during the poly(A)$^+$ RNA extraction (Wroe et al., 2000). The primers used to make the antisense specific cDNA were designed adjacent to the end of the MCRII region [mcr2 G long] and in exon 1 of *Wt1* [WT1 Ex1C]. The primers used to make the sense specific cDNA were designed in the *Wt1* promoter region [MWT ProR] and intron 2 of *Wt1* [WT1 In2.2] (primer positions not shown in Figure 5.1A). Negative control cDNA reactions were also set up with no primers. By using poly(A)$^+$ RNA, rarer transcripts were enriched due to the elimination of non-polyadenylated transcripts; in practice, any transcripts with a run of A-bases will be selected for. Therefore, it was still possible that RT-PCR products could be made with antisense specific cDNA if the antisense transcript was non-polyadenylated, and with sense specific cDNA made using primers designed in regions not present in mRNA (such as the promoter and intronic regions). The RT-PCR reactions were very messy (data not shown) with smears in the RT- lanes suggestive of genomic DNA contamination. However, even with genomic DNA contamination, a product should only be amplified from sense specific cDNA with primers designed in exon 1 and exon 2 of *Wt1* [WT1 Ex1/WT1 Ex2] due to the size of the intervening intronic sequence. The 277bp product was actually present in the following lanes: both RT+ antisense specific lanes; both RT+ sense specific lanes; the RT+ no primer lane. Therefore, once again, endogenous priming was affecting the results of these RT-PCR reactions.

Another way to find the 5' end of *Wt1-AS* would have been to carry out a strand specific Northern hybridization reaction. Strand specific Northern hybridization

protocols differ from standard Northern hybridization protocols (see sections 2.1.16 and 2.1.17), the main differences being: the use of RNA polymerases and rNTPs to make riboprobes, the probe synthesis reaction being carried out in an RNase-free environment and the hybridization reaction being carried out in a formamide-based hybridization buffer at 42°C. Strand specific Northern hybridization was not attempted for this gene as this technique has been tried before without success (see section 3.4.5). The outcome being there was a strong signal from non-specific binding of the 18S and 28S ribosomal units masking any specific binding.

Due to the fact that the RT-PCR approach unexpectedly did not allow us to distinguish between unprocessed sense-derived and true antisense-derived products, it was clear that this was not a reliable method to map the 5' end of the *Wt1-AS* transcript. It was therefore decided to pursue plans to generate a targeting construct with the ability to both knockdown the full length *Wt1* antisense transcript by the positioning of an upstream transcriptional terminator sequence and to eliminate the region of DNA corresponding to the MCRII sequence.

5.4.2 Targeting vector construction

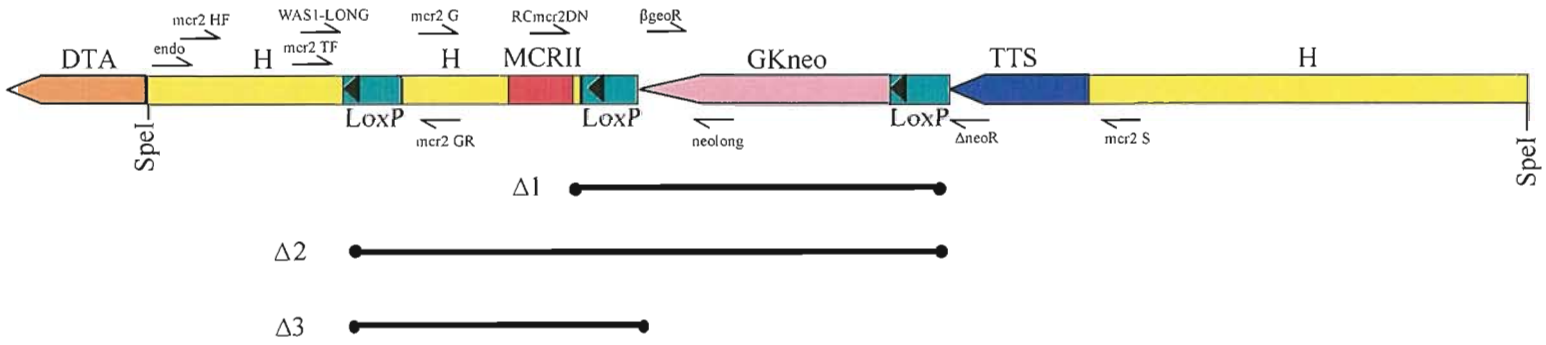
The MCRII parental targeting construct was designed in such a way that *Cre*-loxP deletions would give rise to three secondary alleles: the *Wt1-AS* truncation allele (Δ1), the MCRII targeted deletion (Δ2) and a third unwanted allele (Δ3) (Figure 5.2). The parental targeting construct contained homologous arms to the *Wt1* antisense locus surrounding the MCRII region, three loxP sites, a neomycin positive resistance cassette with its own promoter and polyadenylation signal, a DTA negative selection cassette, the MCRII region and a transcriptional terminator sequence (TTS). The loxP sites were situated so that the neomycin resistance cassette could be removed in both the Δ1 and Δ2 alleles, and the MCRII region could be removed in the Δ2 allele.

Initially, a Wt1 BAC was digested with various restriction endonucleases and a Southern was carried out on these digested DNA fragments using a PCR-derived

Figure 5.2. Schematics showing the design of the targeting constructs generated in order to create mouse models harbouring a truncation of the *Wt1-AS* gene or deletion of the *Wt1-AS* MCRII region plus the truncation. (A) Schematic of the MCRII parental targeting construct used as the basis to generate the two separate mouse models. Shows the positions of each constitutive part of the construct and the *cre*-loxP deletions (black bars beneath the construct) that were required for the generation of the secondary targeting constructs (B) Schematics of the secondary constructs: i, Δ1 results in the premature termination of *Wt1-AS* transcription *in vivo*; ii, Δ2 results in the deletion of the MCRII DNA region *in vivo* as well as the premature termination of *Wt1-AS* transcription; iii, Δ3 is the unwanted allele. Both Δ1 and Δ2 secondary constructs have the neomycin resistance cassette removed. Orange DTA, diphtheria toxin A gene cassette; yellow H, homologous genomic regions; green loxP, loxP recombination sites; red MCRII, MCRII genomic sequence; pink GKneo, neomycin resistance cassette with promoter and polyadenylation signal; blue TTS, transcriptional terminator sequence; arrows, show direction of sequence; ⟶, location of primers.

A

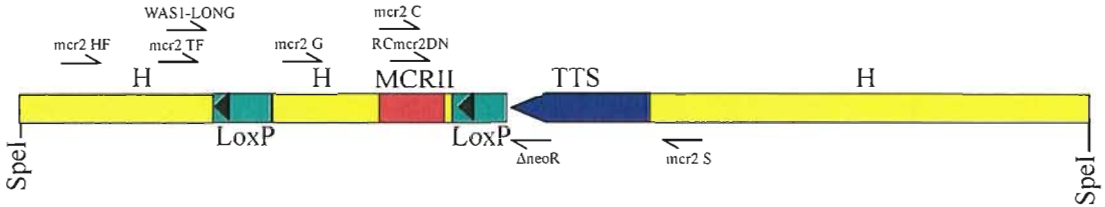
Parental Targeting Construct



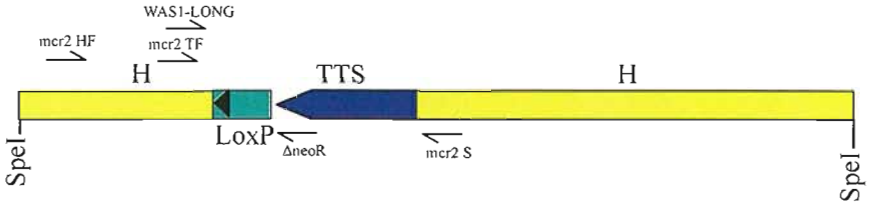
B

Genomic rearrangements expected from loxP recombination

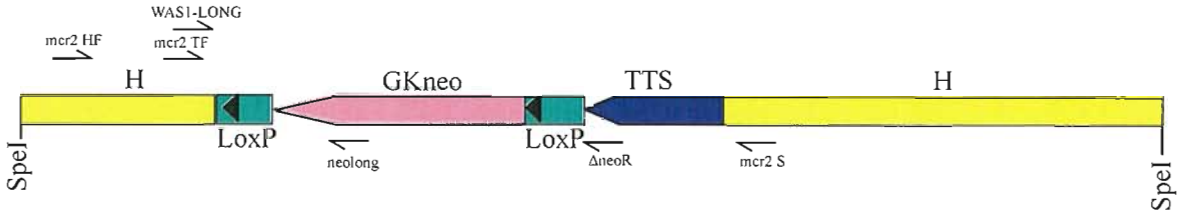
i Δ1: Premature termination of *Wt1 -AS* transcription



ii Δ2: MCRII deletion



iii Δ3: Neomycin resistance cassette intact



probe designed to the MCRII region (work carried out by Dr Andrew Ward, University of Bath). A fragment of a suitable size for cloning was obtained with the enzyme *SpeI*. The Wt1 BAC DNA digested with this enzyme was then sub-cloned into bacterial plasmids using a 'shotgun cloning' approach and colonieswere screened for those containing the correct fragment by Grunstein colony hybridization using the same PCR-derived probe (see section 2.1.14). Sequencing of the ends of the inserts within these sub-cloned plasmids was used to confirm those which contained the required genomic DNA of the *Wt1-AS* gene and which region of the gene had been inserted into each plasmid. The sub-cloned plasmids were digested with numerous different restriction enzymes in order to identifying unique restriction sites that could be used to produce a physical map, thereby providing enough molecular information for subsequent cloning of the targeting construct.

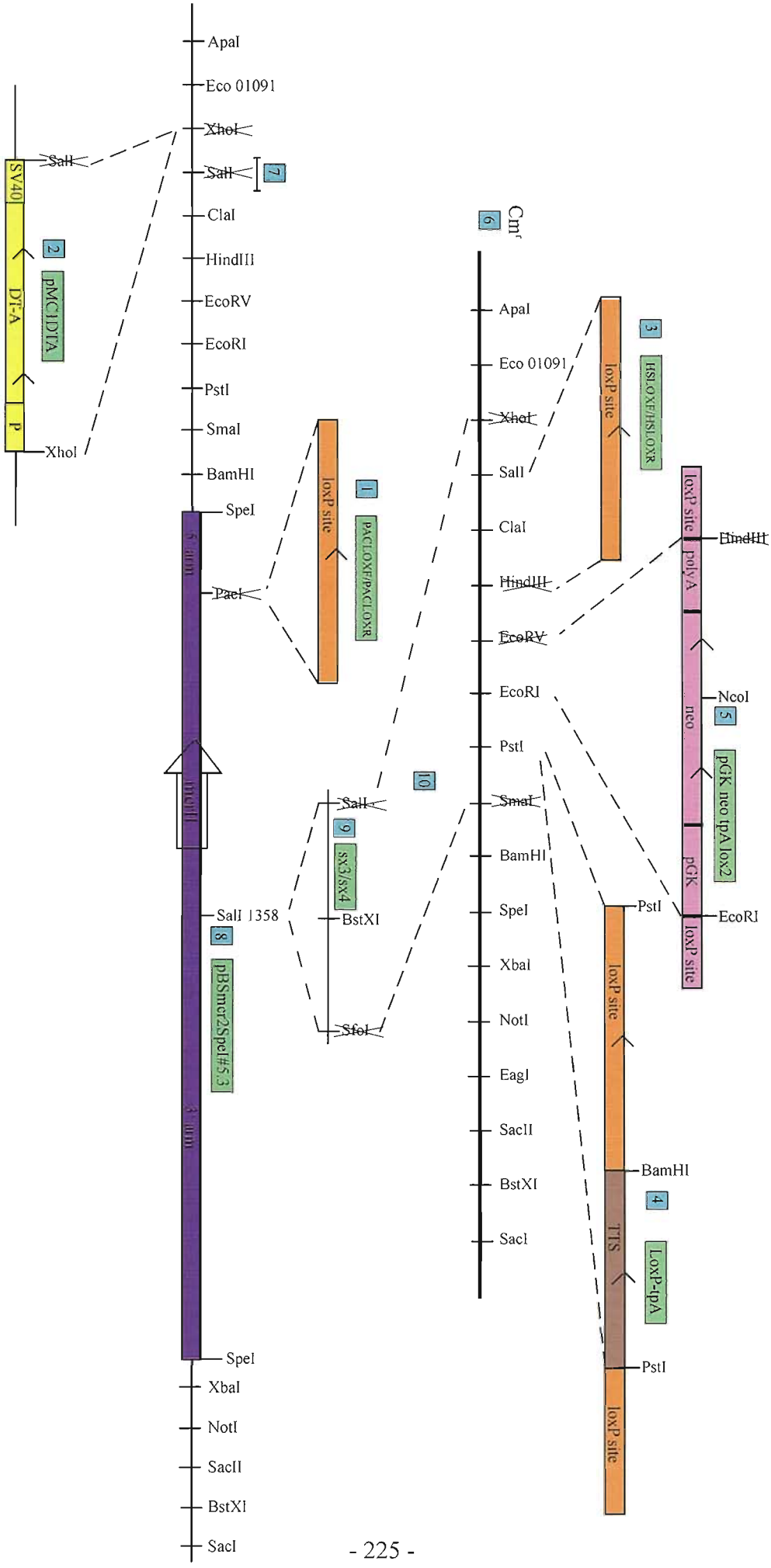
A *SpeI* sub-cloned Wt1 BAC fragment in pBluescript (plasmid pBSmcr2SpeI#5.3), containing part of the mouse *Wt1-AS* gene, including the MCRII region was used as the basis to generate the 5' and 3' homologous arms as well as being the source of the MCRII region. The region of the *Wt1-AS* genomic DNA used for the 5' homologous arm (5' in relation to the *Wt1* gene) was approximately 754bp encompassing the *SpeI* site 5870bp upstream and the *PacI* site 5115bp upstream of the ATG of the *Wt1* gene of Ensemble v39 (entry ENSMUSG00000016458). The region of the *Wt1-AS* genomic DNA used for the 3' homologous arm (3' in relation to the *Wt1* gene) was approximately 3183bp encompassing the GTCAAA site 4506bp upstream and the *SpeI* site 1323bp upstream of the ATG of the *Wt1* gene of Ensemble v39 (entry ENSMUSG00000016458). The MCRII region and surrounding genomic DNA was approximately 595bp encompassing the *PacI* site 5115bp upstream and the GTCAAA site 4506bp upstream of the ATG of the *Wt1* gene of Ensemble v39 (entry ENSMUSG00000016458).

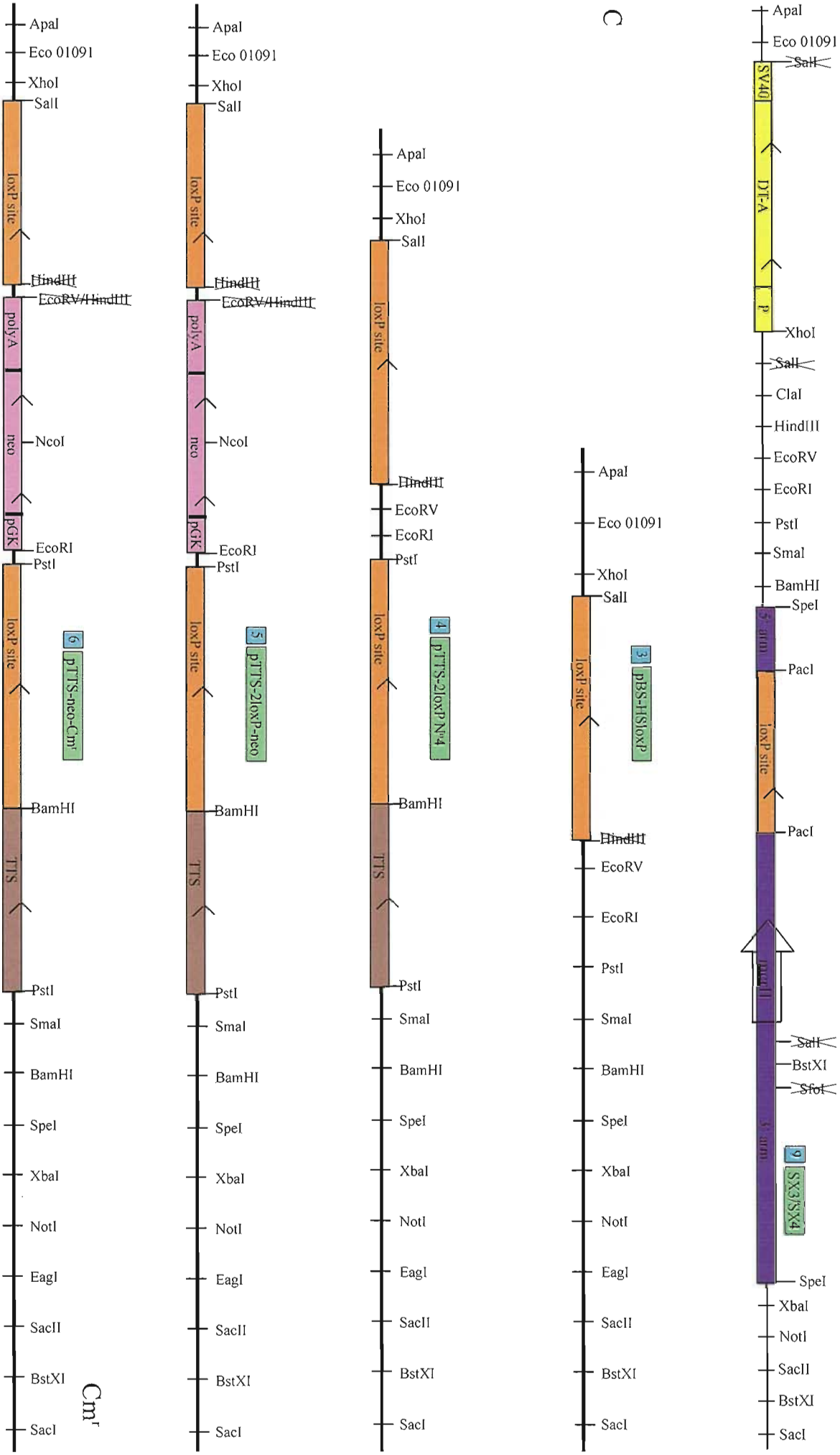
Figure 5.3A shows the steps involved in creating the parental targeting construct. A detailed account of the cloning strategy and each sequential cloning step is discussed below. To check the correct fragments had been inserted into the relevant vectors at each step of the cloning strategy, plasmids were digested to

Figure 5.3. (A) The cloning strategy used to generate the MCRII parental targeting construct: the construct used to generate the MCRII targeted deletion and truncation alleles. Relevant restriction sites are shown within plasmid multiple cloning site regions (thin horizontal lines) and within plasmid inserts (contiguous coloured boxes). Numbered, blue boxes indicate sequential cloning steps. Labelled, green boxes indicate plasmid names. A 4.5kb *SpeI* sub-cloned PAC fragment in pBluescript containing part of the mouse *WT1-AS* gene, including the MCRII region, was used as the basis of the targeting allele that will later form both the 5' and 3' homologous arms (named pBSmcr2SpeI#5.3). 1. An oligonucleotide linker [PACLOXF/PACLOXR] containing the loxP sequence with *PacI* overhangs at either end was inserted into the *PacI* site of pBSmcr2SpeI#5.3 ensuring that the loxP site was cloned in the same orientation relative to MCRII (named pmcr2-loxP). The *PacI* site was not recreated due to the design of the linker. 2. A DTA cassette (from plasmid pMC1DTA) was cloned into pmcr2-loxP. This was achieved by digesting pMC1DTA with *XhoI* and *SalI*, and digesting pmcr2-loxP with *XhoI*. *XhoI* and *SalI* generate compatible sticky-ended cleavage sites and can therefore be directly ligated together, re-generating neither site (named pmcr2-loxP-DTA). 3. An oligonucleotide linker [HSLOXF/HSLOXR] containing the loxP sequence with a *HindIII* overhang at one end and a *SalI* overhang at the other end was inserted into the *HindIII* and *SalI* sites of pBluescript (named pBS-HSloxP). The *HindIII* site was not recreated due to the design of the linker. 4. A TTS cassette and a loxP site (from plasmid loxP-tpA) were cloned into pBS-HSloxP. This was achieved by digesting both pBS-HSloxP and loxP-tpA plasmids with *PstI* and ligating together, ensuring that that TTS and loxP cassettes were cloned in the same orientation relative to the other loxP site (named pTTS-2loxP N°4). 5. A neomycin cassette with a constitutively active promoter and polyadenylation signal (from plasmid pGK-neo-tpA-lox2) was cloned into pTTS-2loxP N°4. This was achieved by digesting pGK-neo-tpA-lox2 with *HindIII* (followed by a fill-in reaction) and *EcoRI* and digesting pTTS-2loxP N°4 with *EcoRV* and *EcoRI*. The blunted *HindIII* site was compatible for re-ligation with the blunt-ended cleavage sites of *EcoRV* (named pTTS-2loxP-neo). 6. In order increase the efficiency of further cloning steps, the ampicillin resistance gene of pTTS-2loxP-neo was replaced by a chloramphenicol resistance gene (named pTTS-neo-Cm^r). This was achieved by recombining a chloramphenicol resistance transposon into the plasmid using the Genome Priming System (NEB). 7. As a unique *SalI* site was required at a later point in the construction, the *SalI* site in pmcr2-loxP-DTA was removed (named pmcr2-DTAΔSalI). 8. A unique *SalI* site was generated upstream of MCRII in pmcr2-DTAΔSalI using the Site Directed Mutageneis Kit (Stratagene) with an oligonucleotide linker [fab1/fab2] containing

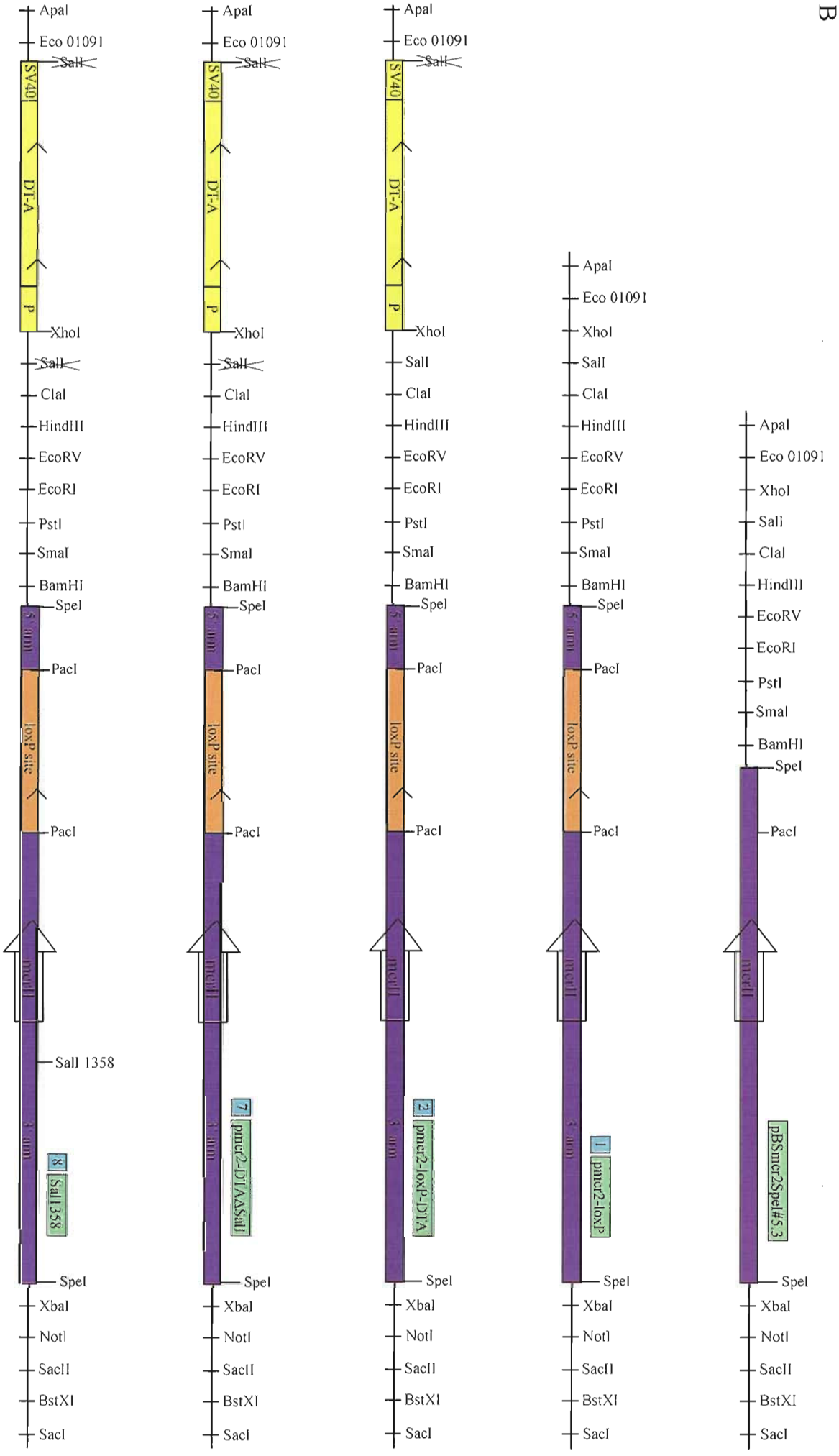
two mismatches (named Sal1358). 9. A blunt ended oligonucleotide linker [sx3/sx4] containing three restriction sites (*SalI*, *BstXI* and *SfoI*) was inserted into the newly formed *SalI* site of Sal1358. This was achieved by digesting Sal1358 with *SalI*, filling-in this site and ligating to the linker (named SX3/SX4). 10. Finally, the two loxP sites, the neomycin resistance cassette and the transcriptional terminator sequence were cloned into SX3/SX4 prior to the start of MCRII using the linker restriction sites. This was achieved by digesting pTTS-neo-Cm^r with *XhoI* and *SmaI* and digesting SX3/SX4 with *SalI* and *SfoI*. *SmaI* and *SfoI* both generate blunt-ended cleavage site and are therefore compatible for re-ligation. *XhoI* and *SalI* have compatible ends and can be ligated together re-generating neither site. (B) The sequential plasmids that made up the left hand region of the targeting construct. (C) The sequential plasmids that made up the right hand region of the targeting construct. (D) The completed construct named mcr2 gtv#2.

A

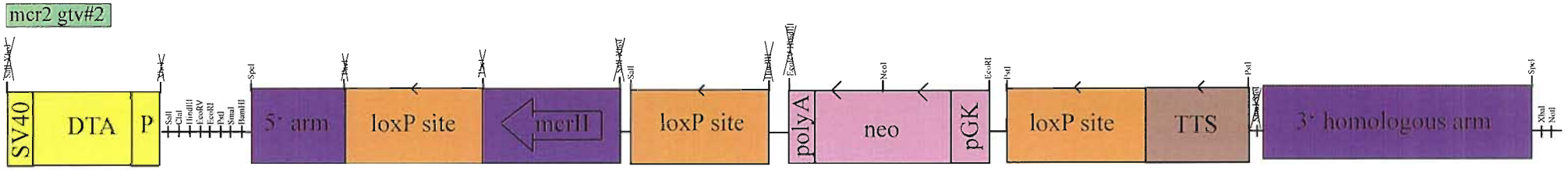

ApaI
Eco 01091
XhoI
SalI
7
ClaI
HindIII
EcoRV
EcoRI
PstI
SmaI
BamHI
SpeI
PacI
5' arm
mcrIII
SalI 1358
8
pBSmcr2SpeI#5.3
3' arm
SpeI
XbaI
NotI
SacII
BstXI
SacI
SalI
SV40
2
pMC1DTA
DT-A
P
XhoI
loxP site
1
PACLOXF/PACLOXR
6
Cmr
ApaI
Eco 01091
XhoI
SalI
ClaI
HindIII
EcoRV
EcoRI
PstI
SmaI
BamHI
SpeI
XbaI
NotI
EagI
SacII
BstXI
SacI
10
SalI
9
sx3/sx4
BstXI
SfoI
loxP site
3
HSLOXF/HSLOXR
loxP site
polyA
HindIII
neo
NcoI
5
pGK neo tpA lox2
pGK
loxP site
EcoRI
PstI
loxP site
BamHI
4
TTS
LoxP-tpA
PstI
loxP site


ApaI
Eco 01091
XhoI
SalI
loxP site
HindIII
EcoRV/HindIII
polyA
neo
NcoI
pGK
EcoRI
PstI
loxP site
6
pTTS-neo-Cmr
BamHI
TTS
PstI
SmaI
BamHI
SpeI
XbaI
NotI
EagI
SacII
BstXI
SacI
Cmr
ApaI
Eco 01091
XhoI
SalI
loxP site
HindIII
EcoRV/HindIII
polyA
neo
NcoI
pGK
EcoRI
PstI
loxP site
5
pTTS-2loxP-neo
BamHI
TTS
PstI
SmaI
BamHI
SpeI
XbaI
NotI
EagI
SacII
BstXI
SacI
ApaI
Eco 01091
XhoI
SalI
loxP site
HindIII
EcoRV
EcoRI
PstI
loxP site
4
pTTS-2loxP N°4
BamHI
TTS
PstI
SmaI
BamHI
SpeI
XbaI
NotI
EagI
SacII
BstXI
SacI
C
ApaI
Eco 01091
XhoI
SalI
loxP site
3
pBS-HSloxP
HindIII
EcoRV
EcoRI
PstI
SmaI
BamHI
SpeI
XbaI
NotI
EagI
SacII
BstXI
SacI
ApaI
Eco 01091
SalI
SV40
DT-A
p
XhoI
SalI
ClaI
HindIII
EcoRV
EcoRI
PstI
SmaI
BamHI
SpeI
5' utr
PacI
loxP site
PacI
Intron II
SalI
BstXI
SfoI
3' utr
9
SX3/SX4
SpeI
XbaI
NotI
SacII
BstXI
SacI

B


ApaI
Eco 01091
SalI
SV40
DT-A
P
XhoI
SalI
ClaI
HindIII
EcoRV
EcoRI
PstI
SmaI
BamHI
SpeI
5' arm
PacI
loxP site
PacI
merI
SalI 1358
3' arm
8
SalI1358
SpeI
XbaI
NotI
SacII
BstXI
SacI
7
pmcr2-DTAASalI
2
pmcr2-loxP-DTA
1
pmcr2-loxP
pBSmcr2SpeI#5.3

D


mcr2 gtv#2
SV40
DTA
P
SalI
ClaI
HindIII
EcoRV
EcoRI
PstI
SmaI
BamHI
SpeI
5' arm
loxP site
mcrII
SalI
loxP site
polyA
NcoI
neo
EcoRI
pGK
PstI
loxP site
TTS
PstI
3' homologous arm
SpeI
XhaI
NotI

look for expected restriction patterns. Following each cloning step, plasmids were sequenced to verify they contained the correct DNA fragments, as any mistakes would be built upon and therefore time consuming to rectify at a later date.

5.4.2.1 Cloning steps

Step 1: A loxP site was inserted into the *PacI* site of the *SpeI* sub-cloned Wt1 BAC fragment in pBluescript by digesting pBSmcr2SpeI#5.3 with *PacI* and annealing with a linker containing the loxP site sequence with *PacI* overhang ends. The two oligonucleotides used to form this linker were PACLOXF/PACLOXR (Table 5.2), the *PacI* site was not recreated due to the design of the linker. Colonies were screened for those that were not digested by *PacI*. Plasmid DNA was then sequenced and selected for the addition of a loxP site cloned in the same orientation relative to MCRII. This construct was named pmcr2-loxP.

Step 2: A DTA cassette (from plasmid pMC1DTA; (Yagi et al., 1990)) was inserted 5' of the 5' homologous arm. pMC1DTA was digested with *SalI* and *XhoI*, pmcr2-loxP was digested with *XhoI* and the fragments ligated together. *XhoI* and *SalI* ends can be directly ligated together, re-generating neither site. Colonies were digested with *XhoI* to detect a larger linear fragment compared to pmcr2-loxP digested with the same enzyme. This construct was named pmcr2-loxP-DTA.

Step 3: A loxP site was cloned into pBluescript. pBluescript was digested with *SalI* and *HindIII* and a linker containing the loxP site sequence with *SalI* and *HindIII* overhang ends was ligated into these sites. The two oligonucleotides used to form this linker were HSLOXF/HSLOXR (Table 5.2), the *HindIII* site was not recreated due to the design of the linker. Colonies were screened for those that were not digested by *HindIII*. This construct was named pBS-HSloxP.

Step 4: A loxP site and a TTS (comprised of a trimerized SV40 polyadenylation signal (Maxwell et al., 1989) from plasmid LoxP-tPA) was next inserted into

pBS-HSloxP. LoxP-tPA and pBS-HSloxP were digested with *PstI* and ligated together. Colonies were screened by looking for a 844bp insert fragment after digestion with *PstI*. Plasmid DNA was then sequenced and selected for the addition of the loxP site and TTS cloned in the same orientation relative to the other loxP site. This construct was named pTTS-2loxP N°4.

Step 5: A neomycin resistance cassette with its own promoter and polyadenylation signal (from plasmid PGK neo tpA lox2, (Soriano, 1999)) was inserted in between the two loxP sites of pTTS-2loxP N°4. PGK neo tpA lox2 was digested with *HindIII* and the overhangs were filled-in with T4 DNA polymerase before digesting with *EcoRI*. pTTS-2loxP N°4 was digested with *EcoRV* and *EcoRI*. *HindIII* once filled-in was compatible with the blunt *EcoRV*, neither site was re-generated following ligation. Colonies were screened by digesting with *EcoRI* which should give a larger linear band compared to pTTS-2loxP N°4 digested with *EcoRI*. This construct was named pTTS-2loxP-neo.

Step 6: In order to increase the efficiency of further cloning steps, the ampicillin resistance gene of pTTS-2loxP-neo was replaced by a chloramphenicol resistance gene. The replacement was achieved by recombining a chloramphenicol resistance transposon into the plasmid using the Genome Priming System (NEB). The kit was designed for the purpose of inserting primer-binding sites randomly within large DNA templates using a Tn7 transposon-based *in vitro* system so that sequencing of the entire template can be carried out. The transposons used in this system contain an antibiotic resistance gene adjacent to the primer-binding site. The kit was advantageous in that it could be used in order to select for colonies that had the chloramphenicol resistance gene inserted at a location that disrupted the ampicillin resistance gene. Only one insertion is possible within each target DNA molecule; therefore the disruption of the ampicillin resistance gene meant that the other regions of the targeting construct remained unaltered. Following transformation of pTTS-2loxP-neo with the GPS kit, colonies were plated onto chloramphenicol agar plates. Only colonies that had integrated the chloramphenicol resistance gene were able to grow on these plates. The colonies were then streaked out onto duplicate agar plates, one with chloramphenicol as before and one with

ampicillin. The colonies that did not grow on the ampicillin plates were those that had integrated the transposon into the ampicillin resistance gene therefore preventing its activity. This construct was named pTTS-neo-Cm^r.

Change in construction design:

The linker HSLOXF/HSLOXR was supposed to also introduce a *SmaI* site next to the *SalI* site; unfortunately for an unknown reason this *SmaI* site was not present in the colony chosen and this problem was not identified at the time. This *SmaI* site along with a second *SmaI* site in pTTS-neo-Cm^r was to have been used to transfer the transcriptional terminator and the neomycin resistance cassette flanked by two loxP sites into an *SspI* site within the genomic DNA making up the 5' and 3' homologous arms. As the *SmaI* site was not present, several further cloning steps had been achieved and the *SspI* site was not unique to pmcr2-loxP-DTA and therefore required a tricky partial digest, it was decided to alter the construction design from this point (Figure 5.3A shows only the final construction design). The new cloning strategy required a unique site within pmcr2-loxP-DTA in which to insert pTTS-neo-Cm^r. As there were no unique sites available that were also present in the correct positions of pTTS-neo-Cm^r, it was decided that a unique site was to be created (step 9) and a linker inserted into this unique site (step 10). The unique site chosen for creation in pmcr2-loxP-DTA was *SalI* (see step 9).

Step 7: As the new cloning strategy required a unique site *SalI* site 3' of the MCRII region, the *SalI* site already in pmcr2-loxP-DTA had to be removed. Removal of the *SalI* site was achieved by its digestion, filling-in and re-ligation to destroy this recognition site. Colonies were screened for those that were not digested by *SalI*. This construct was named pmcr2-DTAΔSalI.

Step 8: The unique site in pmcr2-loxP-DTA was created by site-directed mutagenesis. *SalI* was chosen as there was a 6bp sequence (GTCAAA) in the correct location relative to the MCRII region that only required two base changes to convert it into a *SalI* site (GTCGAC). Site-directed mutagenesis was achieved *in vitro* with a QuikChange® Site Directed Mutagenesis Kit (Stratagene) designed for vector modification. Two complementary purified oligonucleotide

primers [fab1/fab2] were designed to introduce the required modification within the GTCAAA site; these primers were 56bp long containing the modified sequence surrounded by unmodified nucleotide sequence. The kit works in the following way:

- Oligonucleotide primers containing the required modification are allowed to anneal to denatured circular plasmid DNA
- *PfuTurbo* DNA polymerase then replicates each strand of the plasmid incorporating the oligonucleotide primers and forming nicked circular strands
- The template plasmid DNA is then digested with *DpnI*, a restriction endonuclease that only digests methylated DNA
- The unmethylated nicked circular strands containing the modification site are transformed into XL1-Blue supercompetent cells *E.coli* resulting in repair of the nicked DNA

Colonies were screened for those that were digested by *SalI*. This construct was named Sal1358.

Step 9: A blunt ended oligonucleotide linker [sx3/sx4] containing three restriction sites (*SalI*, *BstXI* and *SfoI*) was inserted into the newly formed *SalI* site of Sal1358. The linker was inserted by digesting Sal1358 with *SalI*, filling-in this site and ligating to the linker. Colonies were screened by digesting with *SfoI* that should only give a linerized product if the linker has been inserted. Plasmid DNA was then sequenced to select for the addition of the linker in the correct orientation, with the *SfoI* site furthest away from the MCRII region. This construct was named SX3/SX4.

Step 10: The final step was to transfer the transcriptional terminator and the neomycin resistance cassette flanked by two loxP sites into the genomic DNA making up the 5' and 3' homologous arms prior to the start of the MCRII region. pTTS-neo-Cm^r was digested with *XhoI* and *SmaI*. SX3/SX4 was digested with *SalI* and *SfoI*, two sites that had been added via the linker in the previous step. *XhoI* and *SalI* ends can be directly ligated together, re-generating neither site. *SmaI* and *SfoI* are both blunt ended enzymes and can also be ligated together, re-generating neither site. The efficiency of this cloning step was greatly increased by the different antibiotic resistance genes of the insert and vector plasmids,

preventing any undigested insert plasmids from being retransformed. This final ~11.4kb construct was named mcr2 gtv#2. Figure 5.3B shows the completed construct.

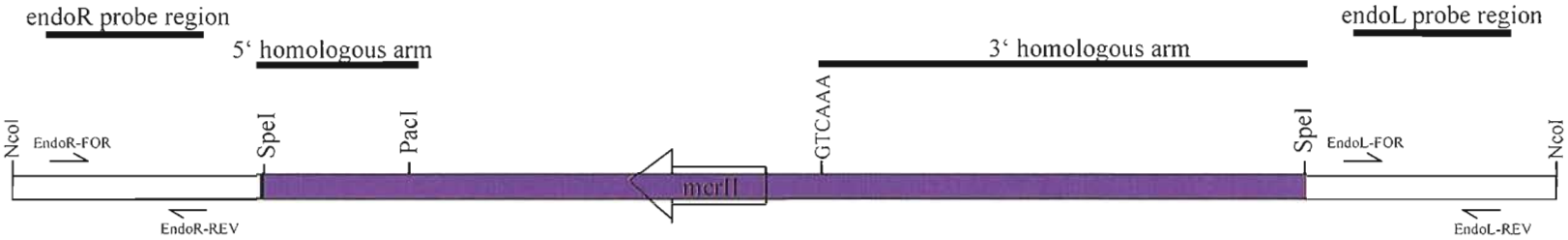
Once cloning was completed, the final construct (mcr2 gtv#2) was linearized at the 3' end, using the unique *NotI* restriction site, and integrated into 129 ES cells following electroporation by homologous recombination of the *Wt1-AS* 3' and 5' homologous regions.

5.4.3 Identification of correctly targeted ES cell clones following the introduction of the targeting construct

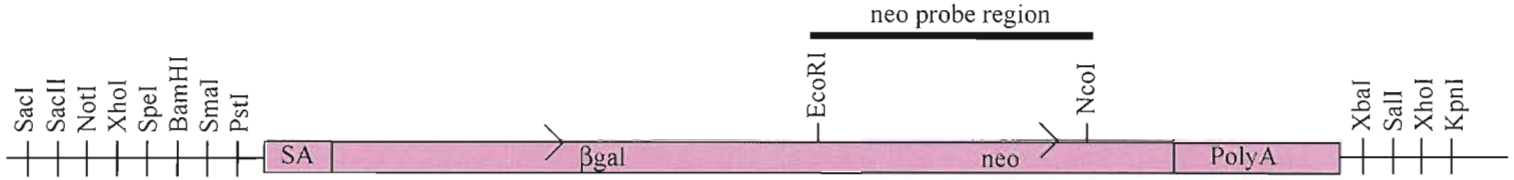
Following transfection, genomic DNA was extracted from 116 ES cell clones that were growing at a similar rate. The genomic DNA was digested with *NcoI* and a Southern hybridization reaction was carried out to determine if the correct targeting had occurred. The probes used were endoL and endoR (Figure 5.4A). Out of these 116 clones, 114 had a clear Southern result with one or other of the probes and one clone seemed to be correctly targeted. The correctly targeted ES cell clone was called M101. As only one clone was positive, the transfection was carried out again. This second time, out of 165 clones, 152 had a clear Southern result with one or other of the probes and four seemed to be correctly targeted. These correctly targeted clones were called C68, C100, C216 and C277. The five correctly targeted clones along with the parental ES cell line (E14C4) and two clones with random integrations (C182 and C260) were thawed and grown so that further genomic DNA could be extracted. The genomic DNA was digested with *NcoI* and Southern hybridization was repeated (Figure 5.4C-E) with probes endoL and endoR to confirm the targeting and an additional probe (neo probe) used to detect the presence of a single neomycin resistance cassette. As expected M101, C68, C100, C216 and C277 were all correctly targeted; they also had only one neomycin resistance cassette as seen by a 8.6kb band on the Southern blot.

Figure 5.4. Schematics of the probes used for Southern blotting and the autoradiograph results of the Southern blotting experiments. A-B show schematic representations of the position of (A) endo L probe and endo R probe in relation to the homologous arms and (B) neo probe in relation to the SAβgeo cassette. (A) The purple box represents the *Wt1* BAC used to generate the 5' and 3' homologous arms; the white box represents the genomic regions upstream and downstream of the homologous arms. (B) The pink box represents the splice acceptor (SA), the β-galactosidase gene (βgal), the neomycin resistance gene (neo) and the polyadenylation signal (PolyA) of the SAβgeo plasmid. C-E show Southern blotting results showing *NcoI* digested DNA hybridisation banding patterns of seven neomycin resistant ES cell clones and the parental cell E14C4 with: (C) endo L probe (D) endo R probe and (E) neo probe. (C) The upper band is the 10.56kb wildtype band, and the lower band is the 8.7kb targeted band. These sizes are very close together and appear as a doublet on the original blot (seen with clones M101, C86, C100, C216, C277). The lanes with the doublet are marked with two dots to the left of the lanes. Correctly targeted clones with the endo L probe are M101, C68, C100, C216 and C277. (D) The upper band is the 10.56kb wildtype band, and the lower band is the 4.8kb targeted band. The targeted bands are marked with a dot to the left of the lanes. As with the endo L probe, correctly targeted clones with the endo R probe are M101, C68, C100, C216 and C277. (E) An 8.6kb band represents clones with a correctly targeted single neo cassette. The targeted bands are marked with a dot to the left of the lanes. As with the endo L and endo R probes, correctly targeted clones are M101, C68, C100, C216 and C277. Clones C182 and C260 are incorrectly targeted.

A


endoR probe region
5‘ homologous arm
3‘ homologous arm
endoL probe region
NcoI
EndoR-FOR
SpeI
PacI
GTCAAA
merII
SpeI
EndoL-FOR
NcoI
EndoR-REV
EndoL-REV

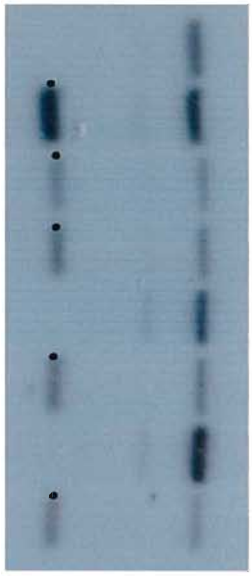
B


neo probe region
SacI
SacII
NotI
XhoI
SpeI
BamHI
SmaI
PstI
EcoRI
NcoI
XbaI
SalI
XhoI
KpnI
SA
βgal
neo
PolyA

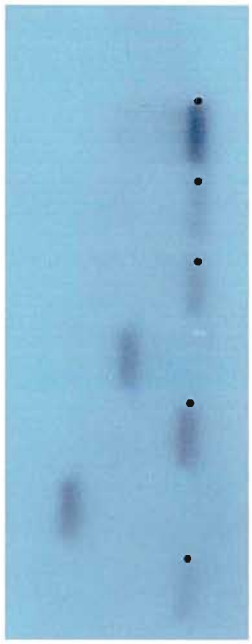
C


E14C4
M101
C68
C100
C182
C216
C260
C277

D


E14C4
M101
C68
C100
C182
C216
C260
C277

E


E14C4
M101
C68
C100
C182
C216
C260
C277

5.4.4 Identification of suitable clones following the introduction of the targeting construct

As only one of the correctly targeted clones (M101, C68, C100, C216, C277) needed to be chosen for the second round of transfection, (where *Cre*-loxP recombination events took place in order to generate the two secondary alleles, Δ1 and Δ2) identification of a suitable clone was undertaken by karyotype analysis. C100 was initially chosen for this analysis as the ES colonies had a nice 'nesty' morphology indicative of a high proportion of undifferentiated cells. Although the number of chromosomes could not be distinguished in many of the cell spreads, in those that could be distinguished, all had the correct complement of 40 chromosomes (Figure 5.5). It was therefore decided to use clone C100 for the second round of transfection with the pCAG-CRE plasmid (Araki et al., 1995) in order to create *Cre*-loxP deletions within the parental targeting construct (see section 5.3.3).

5.4.5 Identification of correctly recombined ES cell clones following transfection of *Cre recombinase*

Genomic DNA was extracted in order to carry out diagnostic PCRs from any of the ES cell clones that were growing at a normal rate and had the appearance of 'nests'; 42 of the picked clones grew sufficiently to extract DNA. PCR reactions with three different primer pairs were then carried out to determine the *Cre*-mediated loxP recombinations within each clone (Figure 5.6A-B). These three sets of primer pairs were designed to amplify between: the MCRII region and the 3' homologous arm [RCmcr2DN/mcr2S] to check for the deletion of MCRII; the MCRII region and the neomycin resistance gene [mcr2 G/neolong] to check for the deletion of the neomycin cassette; and the 5' and 3' homologous arms [mcr2 HF/mcr2 S] to check for the deletion of MCRII and the neomycin cassette (Figure 5.2 shows the position of these primers). Amplification between RCmcr2DN/mcr2S was only possible with Δ1, the parental targeting vector and the wildtype allele (products being of different sizes); amplification between mcr2 G/neolong was only possible with the parental construct; and amplification between mcr2 HF/mcr2 S was possible with Δ1, Δ2, Δ3, the parental targeting vector and the wildtype allele (products being of different sizes). Out of 42

Figure 5.5. Karyotyping of ES cell line C100. (A) The number of chromosomes counted in six spreads for the targeted ES cell line C100. (B) Chromosome spread of the targeted ES cell line C100 stained with DAPI in order to ascertain karyotype counts. Spreads were visualized using fluorescence microscopy with ultraviolet light at 600X magnification with an oil immersion lens.

A

	Chromosome Counts						
C100	40	40	40	40	40	40	40

B

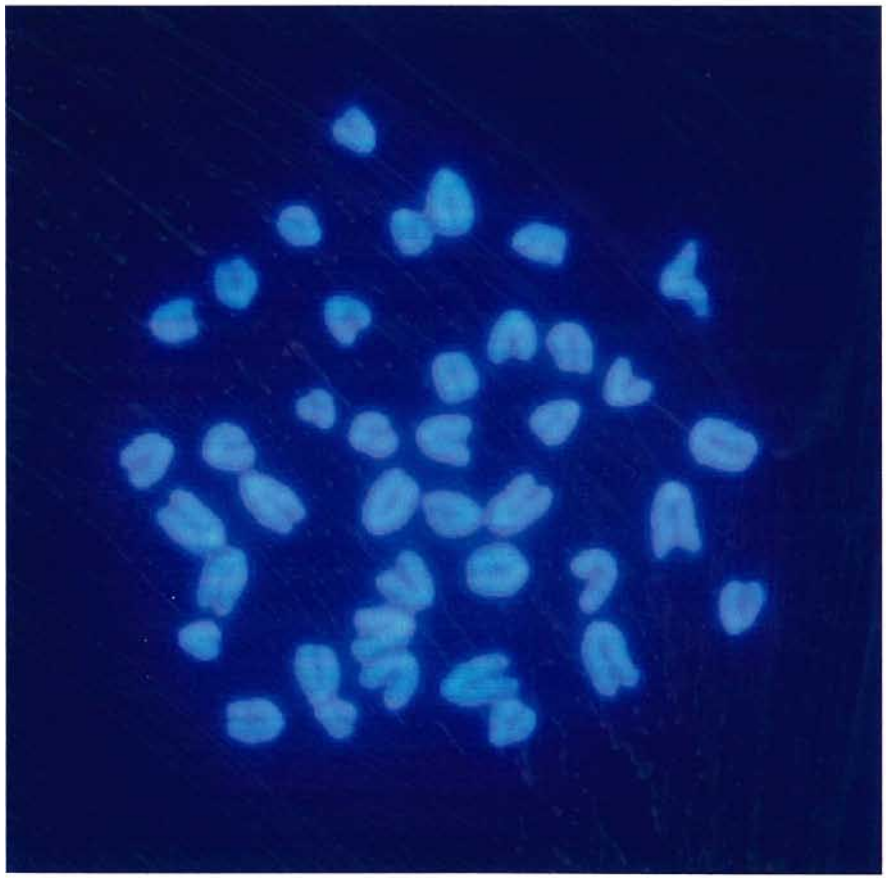
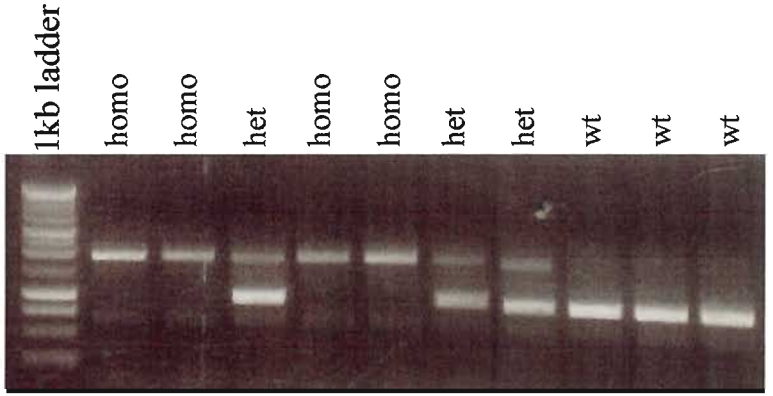


Figure 5.6. PCR to genotype *Cre* recombinase treated ES cell lines. (A) Table showing the PCR size products expected with three different primer pairs for the ES cell clones with a wildtype allele, parental targeting construct and the three deletion constructs. (B) Agarose gels showing PCR products obtained from amplification of ES cell clones with primers RCmcr2DN/mcr2 S (i), mcr2 G/neolong (ii) and mcr2 HF/mcr2 S (iii). i: lanes 1, 3, 4, 5, 7, 9 & 10 show the presence of a 0.2kb and 1.2kb PCR product; lane 6 shows the presence of a 0.2kb and 3kb PCR product; lane 8 shows PCR products from a mixed clone; lane 2 shows no product. ii: lanes 1, 3, 4, 5, 7, 9 & 10 show the absence of a PCR product; lanes 2, 6 & 8 show the presence of a 1kb PCR product. iii: lanes 1, 2, 3, 4 & 6 show the presence of a 1.4kb, 2.4kb and 4.4kb PCR product; lane 5 shows the presence of a 1.4kb, 1.7kb and 4.4kb PCR product. Lanes in i, ii, and iii do not represent the same ES cell clones. (C) Agarose gel showing PCR products obtained from amplification of genomic DNA from ear clips of heterozygous intercross offspring in order to genotype these mice. The 1kb band represents the presence of the wildtype allele and the 2kb band represents the presence of the Δ1 truncation allele. Mice with the 1kb band only are wildtype (wt). Mice with both the 1kb band and the 2kb band are heterozygous (het) for the presence of the wildtype allele and the Δ1 truncation allele. Mice with the 2kb band only are homozygous (homo).

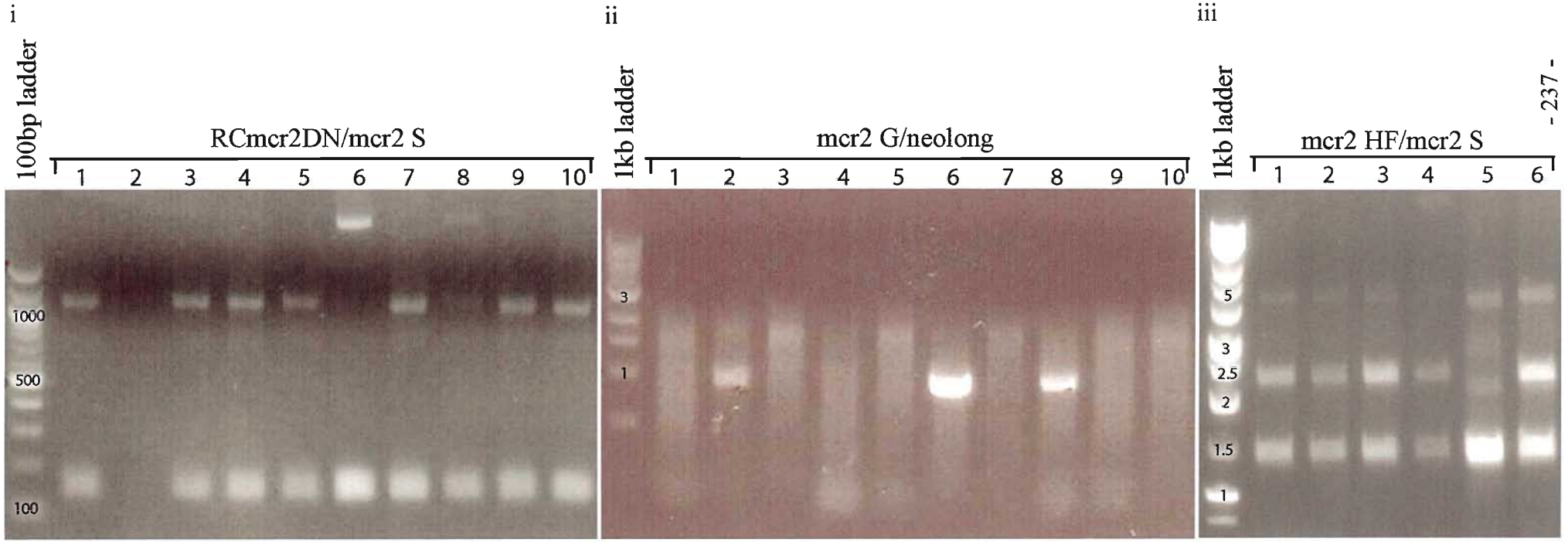
A

	wildtype	parental	Δ1	Δ2	Δ3
RCmcr2DN/mcr2 S	√ (0.2kb)	√ (3kb)	√ (1.2kb)	X	X
mcr2 G/neolong	X	√ (1kb)	X	X	X
mcr2 HF/mcr2 S	√ (1.4kb)	√ (4.4kb)	√ (2.4kb)	√ (1.7kb)	√ (3.8kb)

C



B



clones that grew sufficiently to extract DNA, PCR results were obtained with 26 clones. Out of these 26 clones: 4 had no loxP recombinations; 9 had Δ1; none had Δ2; the rest were mixed clones. As there were no Δ2 clones, the transfection was repeated. This second time, out of the 46 clones that grew sufficiently to extract DNA, PCR results were obtained with 34 clones. Out of these 34 clones: 5 had no loxP recombinations; 4 had Δ1; none had Δ2; the rest were mixed clones. Obviously the process of picking individual clones without antibiotic resistance has lead to the presence of a high number of mixed clones. What is surprising is the absence of any individual Δ2 clones even though there were some mixed clones containing the Δ2 loxP deletion.

5.4.6 Sequencing across the loxP sites of an ES cell clone containing the parental targeting vector

The targeted ES cell line C100 was sequenced across the loxP sites to verify that they were present and the sequences were correct following transfection of the parental targeting construct into cells. Three sets of PCR primer pairs were used to amplify the regions of interest: primers either side of the loxP site between the GKneo cassette and the TTS cassette [βgeoR/mcr2 S], primers between the 5' homologous arm and the GKneo cassette flanking the other two loxP sites [endo/neolong] and primers either side of the loxP site between the 5' homologous arm and the region 5' of MCRII [mcr2 HF/mcr2 GR] (Figure 5.2 for the position of these primers). Following PCR, the products were cloned into pGEM T-easy plasmid (Promega), a plasmid designed for the cloning of PCR products. Plasmid DNA was then extracted and sequenced from a single colony for each primer pair. As a result of these sequencing reactions, it was found that the 5' most loxP site, located between the 5' homologous arm and the MCRII region was absent. The absence of this loxP site provides an explanation as to why there were no Δ2 deletion alleles or Δ3 alleles following transfection of C100 with *Cre* as it was not possible to generate these recombination events. Surprisingly, some of the mixed clones following the second round of transfection seemed, by PCR, to contain both the Δ1 truncation allele and the Δ2 deletion allele. One explanation was that C100 was a mixed ES cell clone with some cells containing the full length parental targeting construct and some cells

containing the parental targeting construct lacking the 5' most loxP site. The other two loxP sites were present and their sequences correct. Subsequently, targeted ES cell lines M101, C68 and C216 were also found to be lacking the 5' most loxP site.

5.4.7 Blastocyst injections

Three of the mixed Δ1/Δ2 clones named cre50, creK20 and creK37 were chosen for blastocyst injections in an attempt to generate chimeric mice able to give rise to offspring containing the Δ1 truncation allele and offspring containing the Δ2 deletion allele. Two chimeric mice (Tiger from creK20 and Travis from creK37) were born from these blastocysts injections (see Table 5.4 for a summary of blastocyst injection efficiency).

5.4.8 Breeding chimeras

The two chimeric mice were bred with wildtype mice of the inbred 129 strain. Coat colour observations could therefore be used to ascertain whether germ line transmission had taken place. An ear clip was taken from chinchilla pups so that DNA could be extracted and the mice genotyped with primers specific for the presence of the wildtype allele (1kb band), the Δ1 truncation allele (2kb band) and the Δ2 deletion allele (1.3kb band) [mcr2 TF/mcr2 S] (see Figure 5.2 for the position of these primers). Both chimeras generated founder lines with transmission of the Δ1 truncation allele through the germline. A second PCR to verify which alleles were present used primers specific for the gene targeting allele (2.6kb), the Δ1 truncation allele (818bp) and the Δ2 deletion allele (153bp) [WAS1-LONG/ΔneoR] (see Figure 5.2 for the position of these primers). The second PCR confirmed that both chimeras generated founder lines with transmission of the Δ1 truncation allele through the germline.

The mcr2 TF/mcr2 S primers were very robust and were therefore also used to screen the offspring of heterozygous intercrosses (Figure 5.6C). The Tiger mice were slow to breed compared to the Travis mice, so the Travis line was used for almost all phenotypic analysis.

Table 5.4

ES cell line/ Parameter	cre50	creK20	creK37
No. transfers	6	2	2
No. blastocysts transferred	54	32	34
No. live births	5	4	6
% transferred blastocysts resulting in live births	9	12	18
No. chimeric offspring	0	1	1
% live births resulting in chimeras	0	25	17
No. chimeras generating founder lines	0	1 Tiger	1 Travis
Overall % transferred blastocysts resulting in live births	**12.5**		
Overall % live births resulting in chimeras	**13.3**		
Overall % of chimeras generating founder lines	**100**		

Table 5.4. Efficiency of blastocyst injections to generate mice with a premature termination of the *Wt1-AS* transcript. Shows the efficiency of obtaining live births, chimeras and founders from blastocyst injections for each targeted cell line used, as well as overall efficiencies.

5.4.9 Sequencing across the loxP sites of ES cell clone creK37

ES cell line creK37 (used to generate the Travis line) was sequenced across the rearranged loxP site to verify that it was present and the sequence was correct following the recombination event. A primer pair [mcr2 C/ΔneoR] was designed to flank this loxP site between the MCRII region and the TTS cassette (see Figure 5.2 for the position of these primers). As before, following PCR, the products were cloned into pGEM T-easy plasmid and DNA was extracted and sequenced from a single colony. The loxP site was intact with the correct sequence.

5.4.10 Study of day 1 pups to check for prenatal death and weight differences between the genotypes

In order to verify that the mice from heterozygous intercrosses were being born in the expected Mendelian ratios, Travis pups were culled at day 1 and a tail clip taken to extract genomic DNA in order to carry out genotyping by PCR. The observed Mendelian ratios were compared to the expected Mendelian ratios using a Chi-squared test in order to determine if the observed values significantly differ from the expected values. Whether differences between the observed and expected values are classed as significantly different from one another depends upon the cut off point for the probability value (*P*); in this study the cut off point was set at $P<0.05$. The results of the Chi-squared test with 23 pups (2 degrees of freedom, $\chi^2=0.391$, $P=0.8223$) indicate that there was no significant difference ($P>0.05$) away from the expected ratios, the conclusion being that pups of all three genotypes were being born as expected (Figure 5.7A). As the number of day 1 pups genotyped was small, 50 mice of weaning age (~3 weeks) were also analysed. The result from this second Chi-squared test (2 degrees of freedom, $\chi^2=0.680$, $P=0.7118$) confirmed that there was no significant difference away from the expected ratios (Figure 5.7A).

The day 1 pups were also weighed and certain organs dissected in order to identify any obvious early differences between the genotypes both by weight and histological analysis. The organs weighed were kidneys, liver, lung and heart. These organ weights have been presented as graphs of weight in grams for each

Figure 5.7. Chi-squared values and graphs representing body and organ weights. (A) Tables of the expected (E) and observed (O) values of wildtype (wt), heterozygote (het) and homozygote (homo) mice born from Travis heterozygous intercrosses. i: day 1 pups. ii: weaning age mice. (B) i-v: Graphs showing body and organ weights in grams of wildtype (wt), heterozygote (het) and homozygote (homo) day 1 pups born from Travis heterozygous intercrosses. i, body weights; ii, kidney weights; iii, liver weights; iv, lung weights; v, heart weights.

A

i

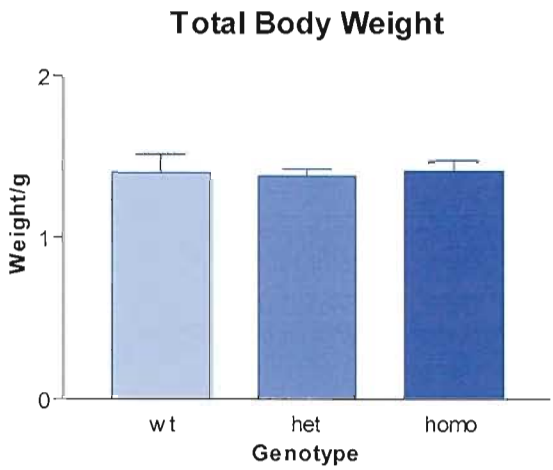
	wt	het	homo
E	5.75	11.5	5.75
O	5	13	5

ii

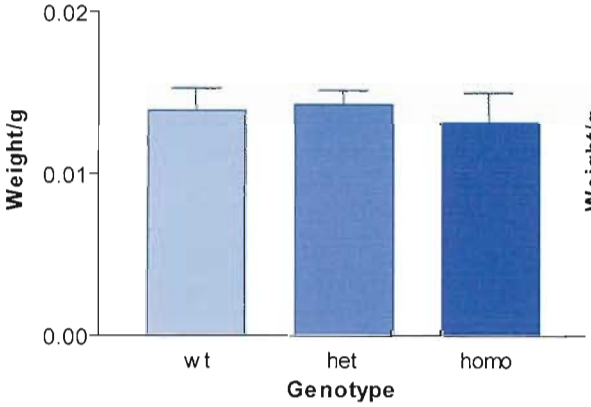
	wt	het	homo
E	12.5	25	12.5
O	15	23	12

B

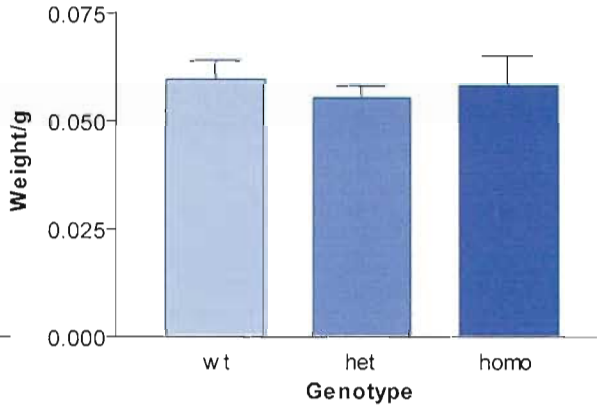
i



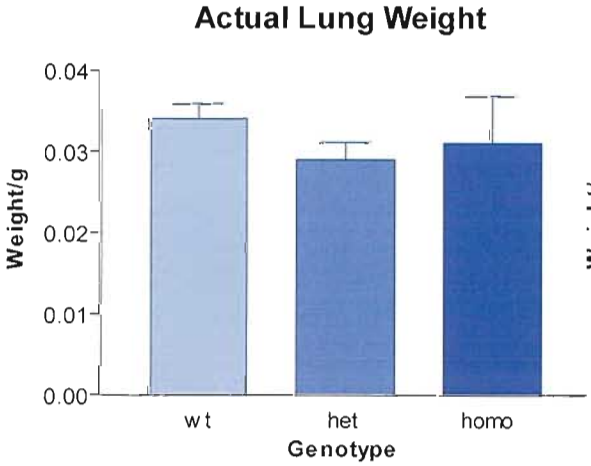
ii



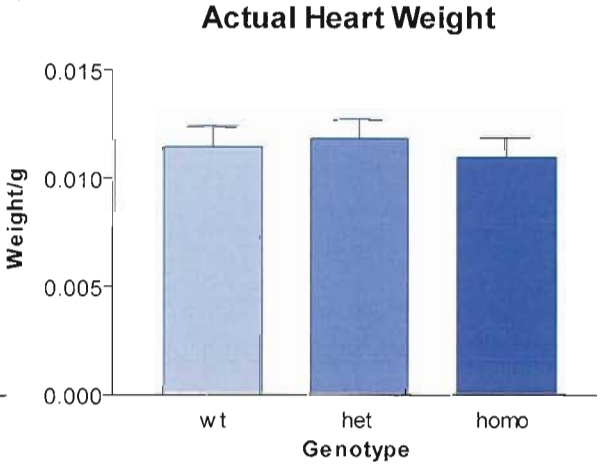
iii



iv



v



genotype (Figure 5.7B) and the values statistically analyzed using one-way analysis of variance (ANOVA) with Tukey's Multiple Comparison post test (Tukey's test). One-way ANOVA is a statistical test used to look for differences between three or more independent groups that follow a Gaussian distribution by giving a *P* value that the groups are sampled from the same population. Whether differences between the three groups were classed as significantly different from one another depends upon the cut off point for the *P* value; in this study the cut off point was set at $P<0.05$. If $P<0.05$, Tukey's post test can be used to compare all groups to each other to find where the differences lay. Statistical tests were performed using GraphPad Prism®. None of the weights were significantly different between the genotypes (Table 5.5A).

In order to detect any obvious histological differences in the kidneys and testes between wildtype and homozygous day 1 pups (Figure 5.8), these organs were fixed, embedded, sectioned and stained (see sections 2.3.1 and 2.3.3). Kidney sections were PAS stained and testis sections were H&E stained (see sections 2.3.5 and 2.3.4). No obvious histological differences were observed; all the appropriate structures could be seen with no difference in the number or distribution of these structures.

5.4.11 Transcriptional termination of the *Wt1-AS* transcript in the Travis line

In order to verify that the transcriptional terminator sequence was causing termination of the *Wt1-AS* transcript in the Travis mice, RT-PCR was carried out on cDNA made from day 2 Travis wildtype or homozygous pups born from a heterozygous intercross in order to detect the presence or absence of *Wt1-AS* (Figure 5.9A-B). Actin specific primers (A11/A12) were used to check that the cDNA had been generated and that there was no genomic DNA contamination (which would have caused RT-PCR products to be made with the cDNA RT- samples). Three gene specific primer pairs were then used: mcr2 CNLong/mcr2 B were designed to amplify the MCRII region and should therefore have only given an RT-PCR product with wildtype pup cDNA; mcr2 SF/mcr2 VR were designed to amplify a region upstream of the transcriptional terminator and

Tables 5.5
A

Day 1 Weights	**Actual**
Body	P=0.9265
Kidneys	P=0.8153
Liver	P=0.7093
Lung	P=0.4890
Heart	P=0.8443

B

7-8 Month Old Weights	**Males**		**Females**	
	Actual	**Relative**	**Actual**	**Relative**
Body	P=0.3634	N/A	**P=0.0187**	N/A
Left Kidney	P=0.2105	P=0.4186	P=0.5403	P=0.3935
Right Kidney	P=0.0676	P=0.2760	P=0.3002	P=0.2823
Left Testis	P=0.4464	P=0.4069	N/A	N/A
Right Testis	P=0.5003	P=0.4098	N/A	N/A
Spleen	P=0.6831	P=0.6871	P=0.9145	P=0.8533
Liver	P=0.5444	P=0.2408	P=0.1055	P=0.4863
Heart	P=0.9528	P=0.8486	P=0.0764	P=0.1532
Thymus	P=0.4072	P=0.5259	P=0.1380	P=0.2191
Brain	P=0.5039	P=0.6656	P=0.4348	**P=0.0247**
Eyes	**P=0.0131**	P=0.7262	P=0.1181	**P=0.0395**

C

1 Year Old Weights

	Males	
	Actual	**Relative**
Body	P=0.7956	N/A
Left Kidney	P=0.6274	P=0.4117
Right Kidney	P=0.0987	P=0.1714
Left Testis	**P=0.0051**	**P=0.0306**
Right Testis	**P=0.0461**	**P=0.0263**
Spleen	P=0.1400	P=0.3355
Liver	P=0.3354	P=0.3673
Heart	P=0.2984	P=0.2443
Thymus	P=0.2003	P=0.1662
Brain	P=0.8004	P=0.8374
Eyes	**P=0.0039**	P=0.3783

Tables 5.5. *P* values of statistical analyzes generated from actual weight data and where appropriate, relative weight data for each of the organs dissected from wildtype, heterozygous and homozygous Travis mice. Table A: One-way ANOVA *P* values for organ weights dissected at day 1. Table B: One-way ANOVA *P* values organ weights dissected at 7-8 months. Table C: T-test *P* values for organ weights dissected at 1 year.

Figure 5.8. Kidney and testis histology. (A-B) Day 1 PAS stained kidney sections (40x). A, Travis wildtype. B, Travis homozygote. (C-D) Day 1 haematoxylin and eosin stained testis sections (200x). C, Travis wildtype. D, Tiger homozygote. d, distal convoluted tubule; p, proximal convoluted tubule; c, comma-shaped body; s, s-shaped body; g, glomerulus; a, interlobular artery; sf, seminiferous tubule; l, leydig cells.

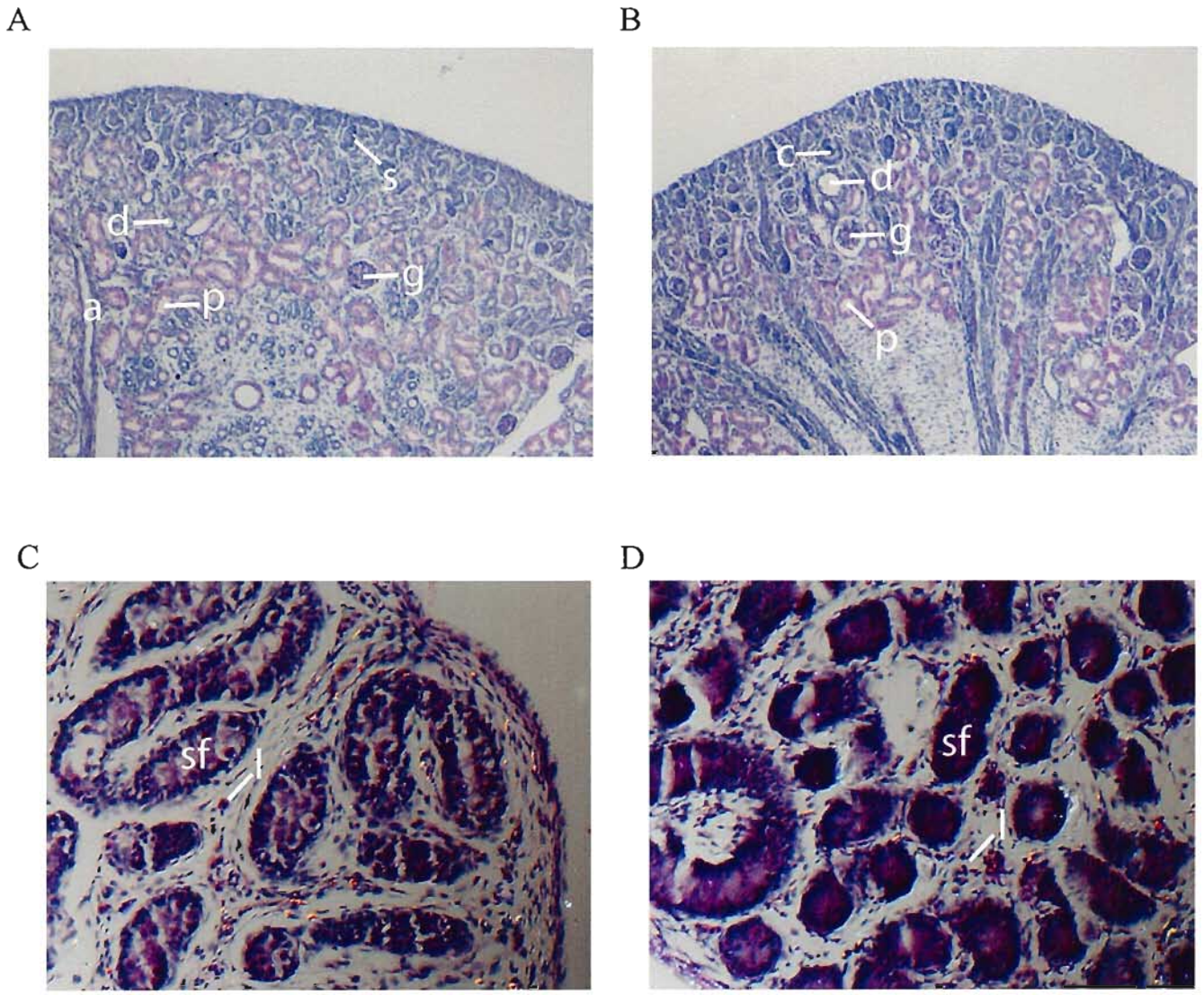
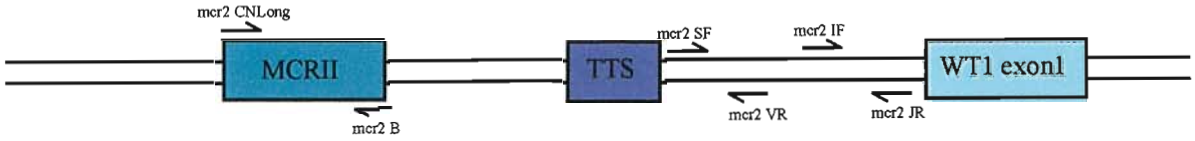
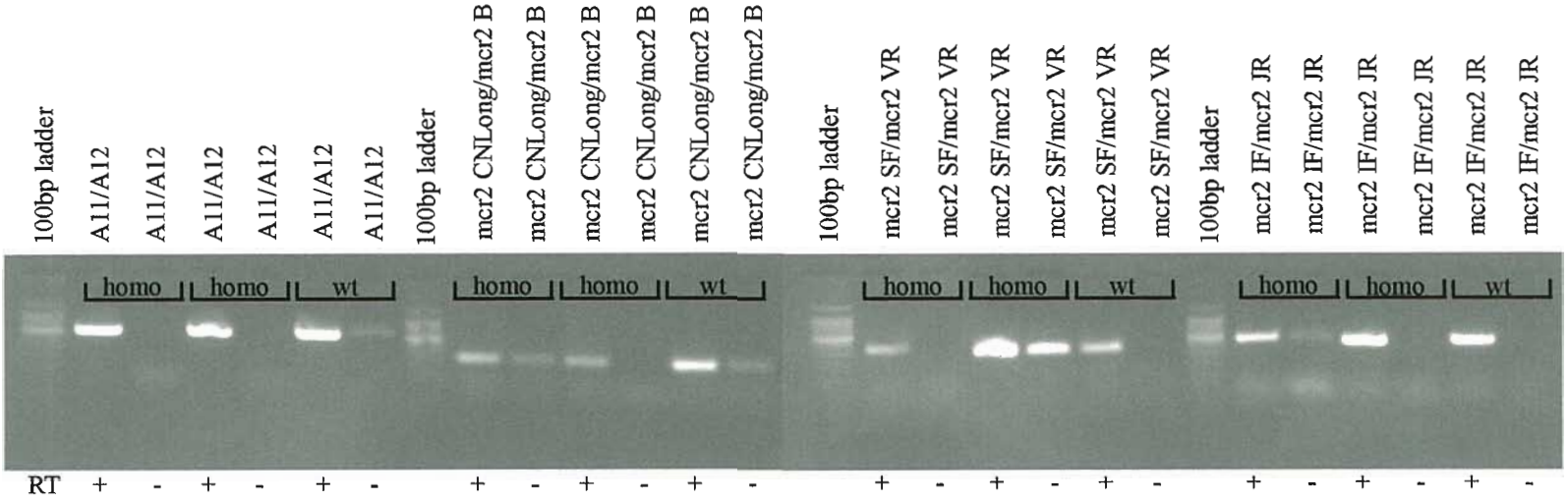

A
s
d
g
a
p
B
c
d
g
p
C
sf
l
D
sf
l

Figure 5.9. RT-PCR to establish if the transcriptional terminator was working within Travis homozygous mice. (A) Schematic of the 5' Wt1 locus showing the MCRII region of the Wt1 antisense transcript, the transcriptional terminator (TTS) of the targeting construct, WT1 exon 1 and the relative position of the primer pairs designed around this locus. (B) Agarose gel showing RT-PCR products obtained with cDNA made from day 2 Travis wildtype (wt) or Travis homozygous (homo) pups using primers A11/A12 (actin), mcr2 CNLong/mcr2 B (designed within the mouse MCRII region), mcr2 SF/mcr2 VR (designed upstream of the transcriptional terminator), mcr2 IF/mcr2 JR (designed further upstream of the transcriptional terminator). Lanes represent RT-PCR reactions carried out with first strand synthesis products made with (RT +) or without (RT -) reverse transcriptase.

A


mcr2 CNLong
MCRII
TTS
mcr2 SF
mcr2 IF
WT1 exon1
mcr2 B
mcr2 VR
mcr2 JR

B


100bp ladder
A11/A12
A11/A12
A11/A12
A11/A12
A11/A12
A11/A12
100bp ladder
mcr2 CNLong/mcr2 B
mcr2 CNLong/mcr2 B
mcr2 CNLong/mcr2 B
mcr2 CNLong/mcr2 B
mcr2 CNLong/mcr2 B
mcr2 CNLong/mcr2 B
100bp ladder
mcr2 SF/mcr2 VR
mcr2 SF/mcr2 VR
mcr2 SF/mcr2 VR
mcr2 SF/mcr2 VR
mcr2 SF/mcr2 VR
mcr2 SF/mcr2 VR
100bp ladder
mcr2 IF/mcr2 JR
mcr2 IF/mcr2 JR
mcr2 IF/mcr2 JR
mcr2 IF/mcr2 JR
mcr2 IF/mcr2 JR
mcr2 IF/mcr2 JR
homo
homo
wt
homo
homo
wt
homo
homo
wt
homo
homo
wt
RT
+ - + - + - + - + - + - + - + - + - + - + - + - + -

should therefore have given an RT-PCR product with both the wildtype and homozygous pup cDNA; mcr2 IF/mcr2 JR were designed further upstream of the transcriptional terminator, and so should have also given an RT-PCR product with both wildtype and homozygous pup cDNA. For each sample there was a product in the RT- lane for at least one of the primer pairs, even though the cDNAs had been treated with DNases to get rid of contaminating genomic DNA. The results were considered to be unreliable as any of the bands could have been amplified from genomic DNA. The cDNA synthesis reactions and the RT-PCRs were repeated several times but the bands amplified varied each time. Working out whether the transcriptional terminator was working by using RT-PCR was therefore abandoned.

Experiments are currently being carried out by Dr Kim Moorwood and Thuluz Meza Menchaca (University of Bath) and Dr Anne Hancock (CLIC Sargent Research Unit, University of Bristol) in order to determine if the transcriptional terminator is working; experiments at University of Bath involve RPA and experiments at University of Bristol involve quantitative PCR (qPCR).

5.4.12 Travis cohort study to check for phenotypic differences between the genotypes

Once a large enough cohort had been generated by breeding, tests were carried out to look for phenotypic differences between wildtype, heterozygous and homozygous siblings. The tests carried out were: urine analysis to look for the presence of protein, an early indication of kidney disease; a weight study of organs that usually express *Wt1*; histological analysis of the same organs.

5.4.12.1 Urine study

Urine was collected (see section 2.4.8) at three different time points: 4 months, 7-8months and 1 year of age. The largest cohort of samples were collected at the first two time points where 60 mice in total were tested for the presence of protein in the urine: 20 wildtype, 20 heterozygote and 20 homozygote siblings from a Travis heterozygous intercross (with an even number of males and

females). The 1 year cohort was much smaller with 12 mice in total: 6 wildtype 129 males and 6 homozygote males from a Travis heterozygous intercross. The urine was run on an SDS-PAGE gel followed by staining with coomassie brilliant blue (see section 2.1.18). It was sometimes difficult to decide when a sample was positive, with protein levels above background, or negative. The decision was therefore a subjective evaluation reached by comparing the intensities between the samples within each gel. As can be seen from the tables of results for protein positive and negative urine samples (Figure 5.10A), only one male homozygote had the presence of protein in the urine at 4 months, whereas at 7-8 months these numbers had increased to 9 males and 3 females with protein in the urine; the mice were of all three genotypes, however, so the occurrence of proteinuria could not be put down to genetic differences between the mice. At 1 year, 1 wildtype and 3 homozygotes has the presence of protein in the urine. Again, as both genotypes were affected and there was no difference in the amount of protein present between the genotypes, the occurrence of proteinuria could not be put down to genetic differences between the mice.

5.4.12.2 Weight study

Weight studies were carried out at two different time points. The first weight study was carried out at 7-8 months of age with 60 mice in total: 20 wildtype, 20 heterozygote and 20 homozygote siblings from a Travis heterozygous intercross (with an even number of males and females). The second weight study was carried out at 1 year of age with 6 wildtype 129 males and 6 homozygote males from a Travis heterozygous intercross. The organs weighed were kidneys, testes, spleen, liver, heart, thymus, brain and eyes. Any mice that exhibited morbid features such as infections or tumours (Table 5.6) were excluded from the weight study and replaced, if available, by another mouse of the same sex and genotype. The weights were presented in two graphical forms: weight in grams and relative weights (weight as a percentage of total body weight). Males and females were separated during the statistical analysis in order to identify any effects of sex. These values were statistically analyzed using one-way analysis of variance (ANOVA) with Tukey's multiple comparison post test (Tukey's test) for the 7-8 month old cohort where there were three groups (Table 5.5B), and using a two-

Figure 5.10. Urine analysis for early detection of proteinuria in Travis mice. (A) Tables showing the number of male and female wildtype (wt), heterozygote (het) and homozygote (homo) mice from a Travis heterozygous intercross that were found to have protein in their urine. i, 4-5 month old mice; ii, 7-8 month old mice; iii, 1 year old mice. (B) SDS-PAGE gels showing the presence or absence of protein in the urine of wildtype (wt), heterozygote (het) and homozygote (homo) mice from a Travis heterozygous intercross. i, 4-5 month old mice; ii, 7-8 month old mice; iii, 1 year old mice. BSA is a positive control for the detection of protein. - lanes are empty. ii: lanes 6 and 11 are positive; iii: lanes 1, 2, 6 and 8 are positive.

A

i

4-5 month urine results			
Males	**wt**	**het**	**homo**
+	0	0	1
-	10	10	9
Females			
+	0	0	0
-	10	10	8

ii

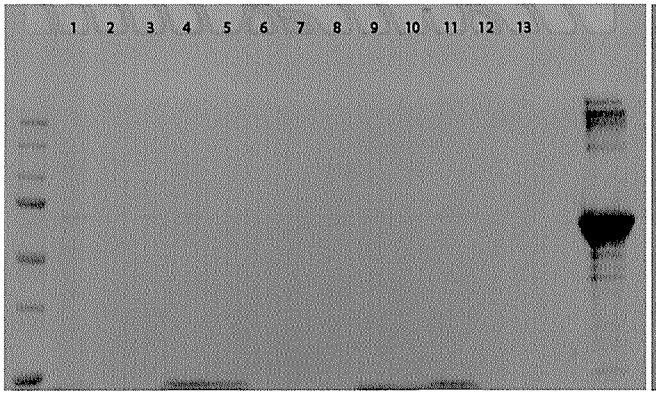
7-8 month urine results			
Males	**wt**	**het**	**homo**
+	2	5	2
-	8	6	7
Females			
+	1	2	0
-	9	8	8

iii

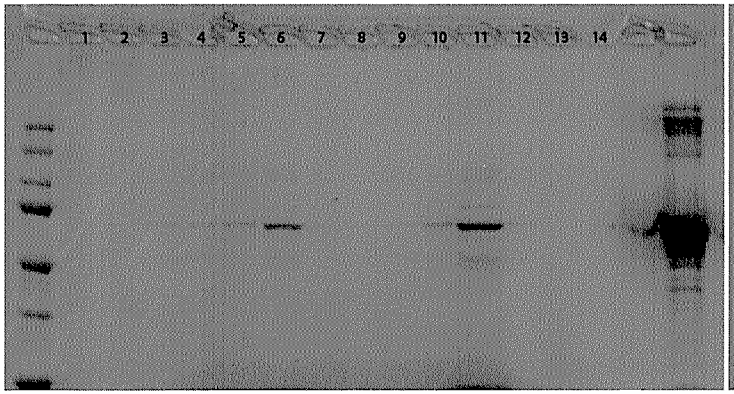
1 year urine results		
Males	**wt**	**homo**
+	1	3
-	5	3

B

i



ii



iii

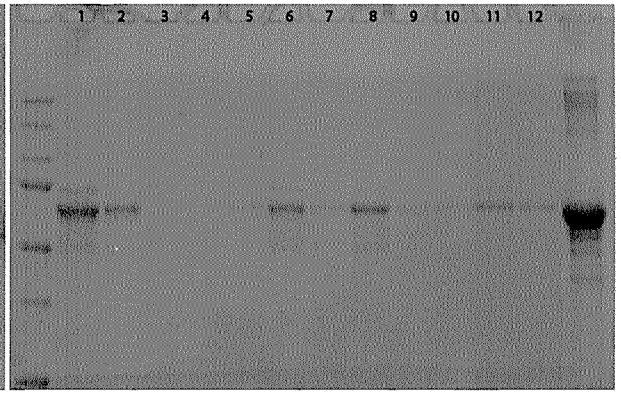


Table 5.6

Genotype	Sex	ill and culled or found dead	Age in months	Reason
wt	♂	dead	1.5	unknown
het	♂	culled	2	eye infection and abnormal growth between the skin and body wall near testes
het	♀	culled	3	large mammary gland abscess and smaller abnormal growths throughout body
het	♂	culled	3.5	tumour surrounding the eye
het	♀	culled	3.5	appeared ill and thin
het	♂	culled	4	tumour surrounding the eye
het	♀	culled	6.5	obese
het	♀	culled	7	obese
het	♀	culled	7.5	appeared ill and thin
wt	♀	culled	7-8	initially part of cohort: liver riddled with tumours (one larger, two smaller)
het	♂	culled	7-8	initially part of cohort: hard white fatty growth in liver that looked like a tumour
homo	♂	culled	8	loss of motor coordination: stumbling, turning in circles, head shaking, seemingly unable to remain still
homo	♀	culled	8	loss of motor coordination: stumbling, turning in circles, head shaking, seemingly unable to remain still

Table 5.6. Travis mice with morbid features. Lists the genotype, sex, if they were culled or were found dead, their age in months, the reason they were culled or were found dead if this information was known.

tailed t-test with the 1 year cohort where there were two groups (Table 5.5C). A t-test is used to look for differences between the means of two independent groups that follow a Gaussian distribution by giving a *P* value that the groups are sampled from the same population. The cut off *P* value for the t-test in this study was set at $P<0.05$.

- 7-8 month study

Body

Firstly, total body weights were analysed (Figure 5.11A). $P=0.3634$ for male body weight and $P=0.0187$ for female body weight. Therefore there was a significant difference between the female body weights. Tukey's post test shows that $P<0.05$ between wildtypes (wt) and heterozygotes (het), and also between het and homozygotes (homo). Het being heavier that both wt and homo. There was no significant difference between the male body weights.

Kidneys

As the left and right kidneys are known to differ in weight, the two kidneys were weighed and analysed separately. Graphs showing kidney weights can be seen in Figure 5.11B i-iv, and showing relative kidney weights can be seen in Figure 5.11C i-iv. There were no significant differences between the groups.

Testes

Left and right testes were weighed and analysed separately. Graphs showing testis weights can be seen in Figure 5.11B v-vi, and showing relative testis weights can be seen in Figure 5.11C v-vi. There were no significant differences between the groups.

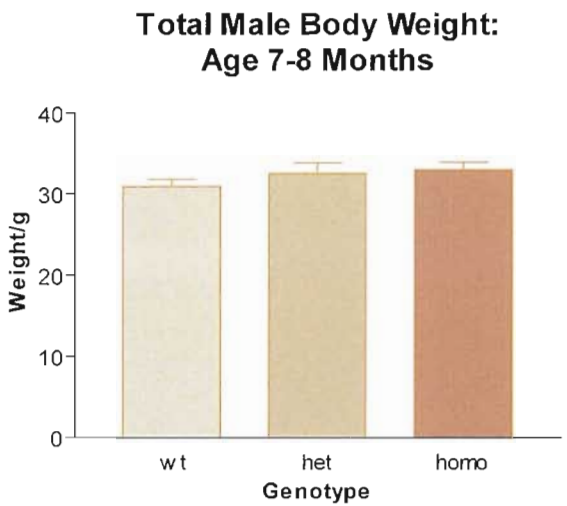
Spleen

Graphs showing spleen weights can be seen in Figure 5.11B vii-viii, and showing relative spleen weights can be seen in Figure 5.11C vii-viii. There were no significant differences between the groups of females. The Bartlett's test for males however had a *P* value <0.05. Bartlett's test is a test for equal variance (standard deviation squared), and since ANOVA is based on the assumption that the populations all have the same variance, the raw data can not be used. In order to equalise the variances, the data was transformed by converting the weight data to reciprocal values; ANOVA was then carried out again with these new values. There were no significant differences between the groups of males in this second ANOVA.

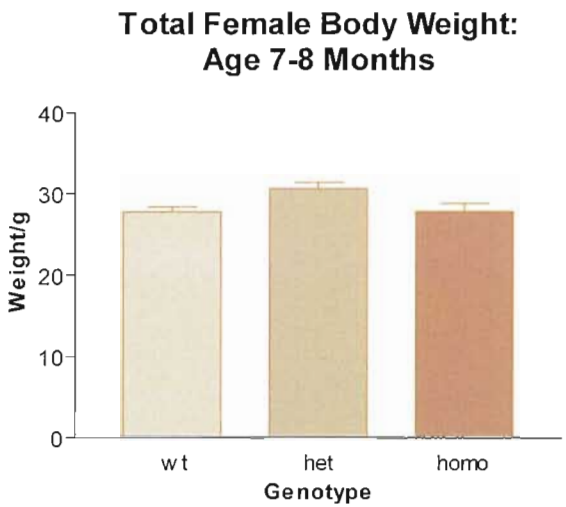
Figure 5.11. Graphs representing body and organ weights of 7-8 month old mice. (A-B) Graphs showing body and organ weights in grams. (C) Graphs showing relative organ weights (organ weight as a percentage of total body weight). The sample populations are wildtype (wt), heterozygote (het) and homozygote (homo) 7-8 month old mice born from a Travis heterozygous intercross. (A) Graphs showing body weights in grams. i, males; ii, females. (B) i-xviii: Graphs showing organ weights in grams. i, male left kidney; ii, female left kidney; iii, male right kidney; iv, female right kidney; v, left testis; vi, right testis; vii, male spleen; viii, female spleen; ix, male liver; x, female liver; xi, male heart; xii, female heart; xiii, male thymus; xiv, female thymus; xv, male brain; xvi, female brain; xvii, male eyes; xviii, female eyes. (C) i-xviii: Graphs showing relative organ weights (graphs in the same order as in B).

A

i

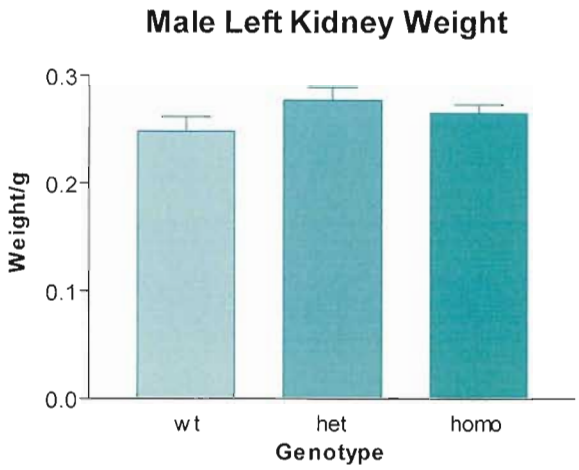

Total Male Body Weight:
Age 7-8 Months
Weight/g
40
30
20
10
0
wt
het
homo
Genotype

ii

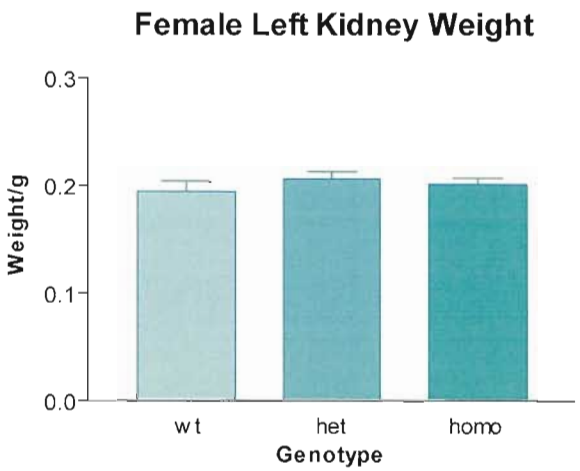

Total Female Body Weight:
Age 7-8 Months
Weight/g
40
30
20
10
0
wt
het
homo
Genotype

B

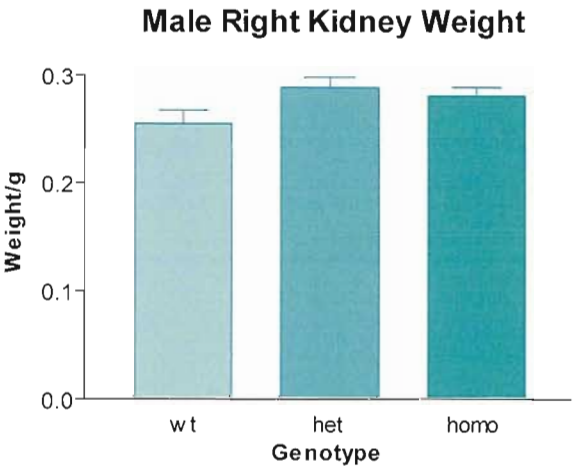
i



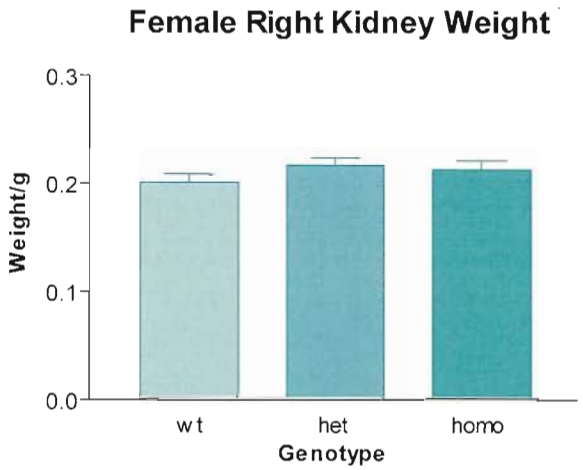
ii



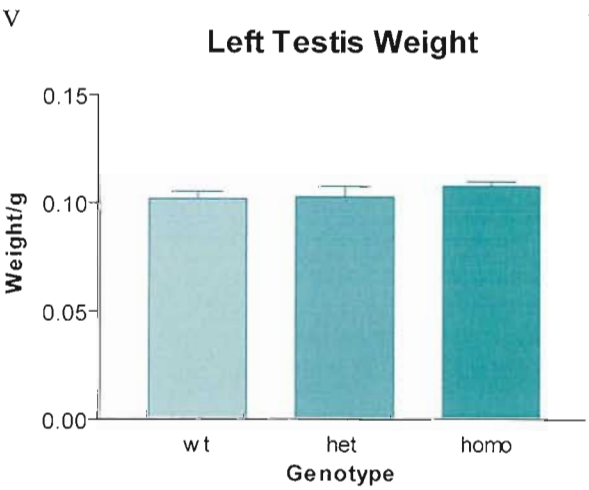
iii



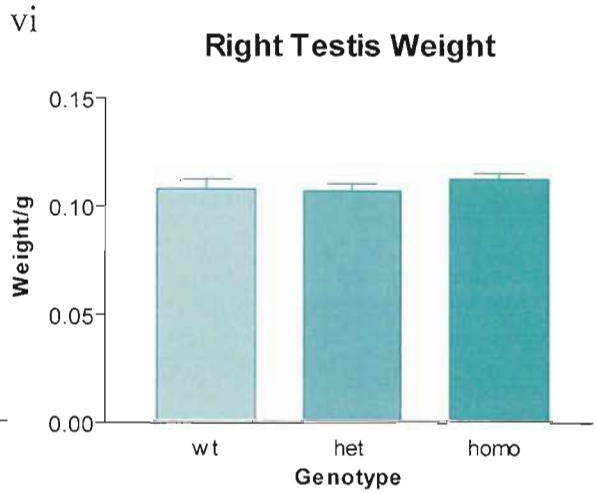
iv



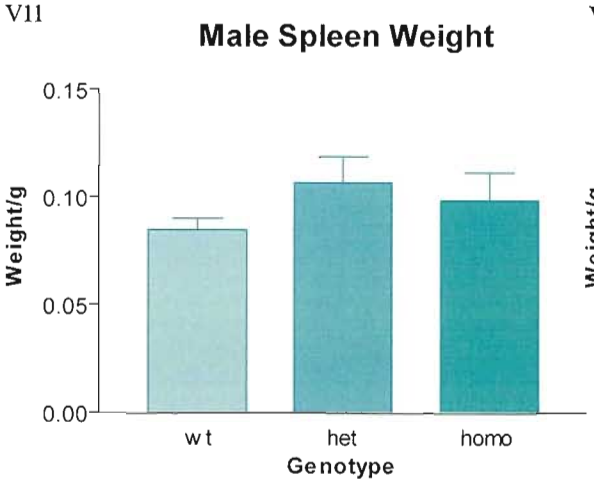
v



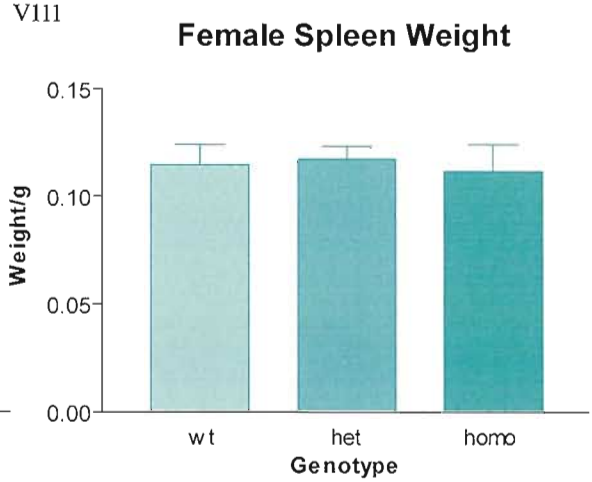
vi



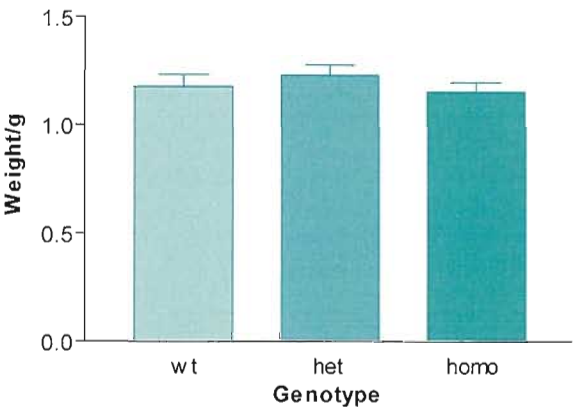
vii



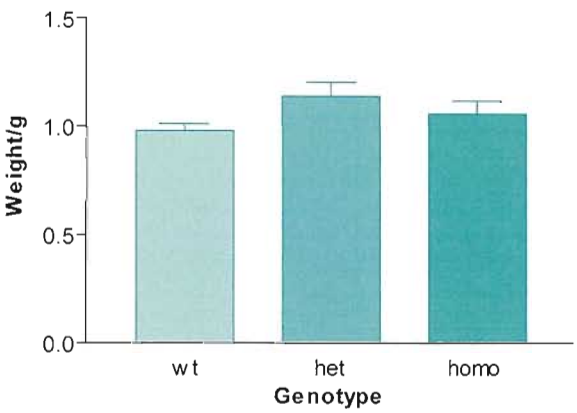
viii



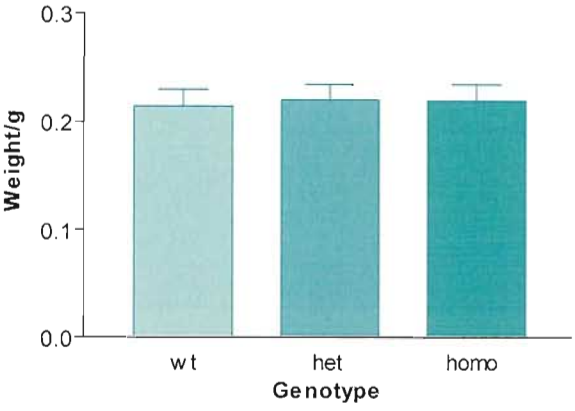
ix



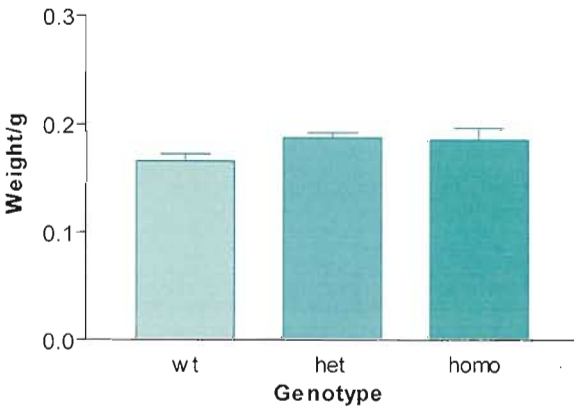
x



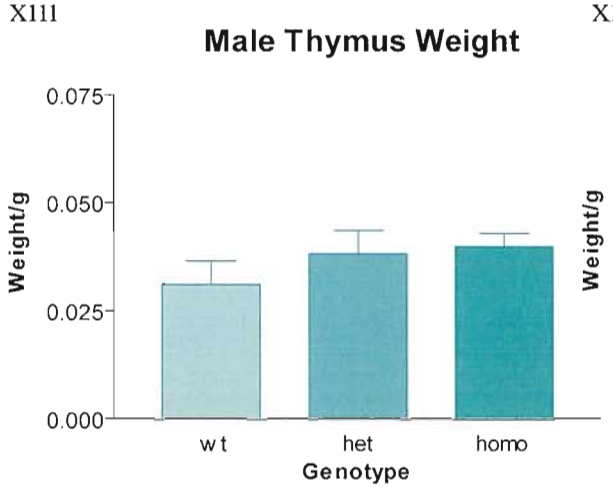
xi



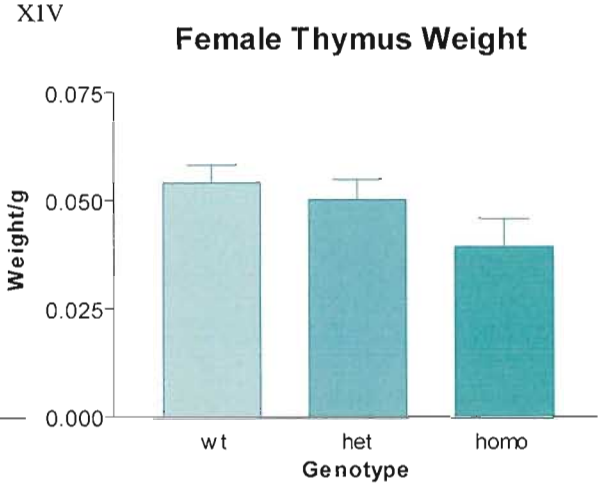
xii



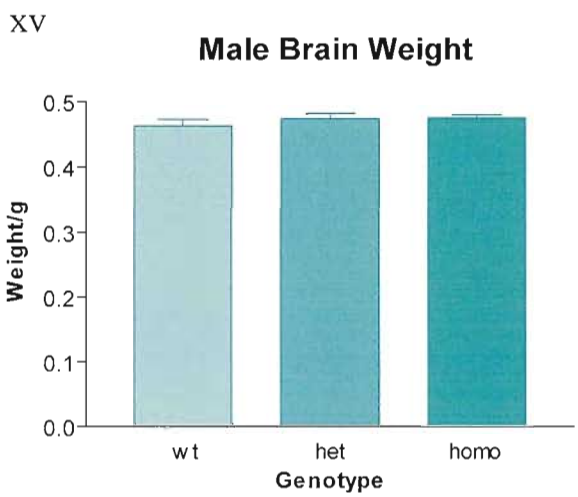
xiii



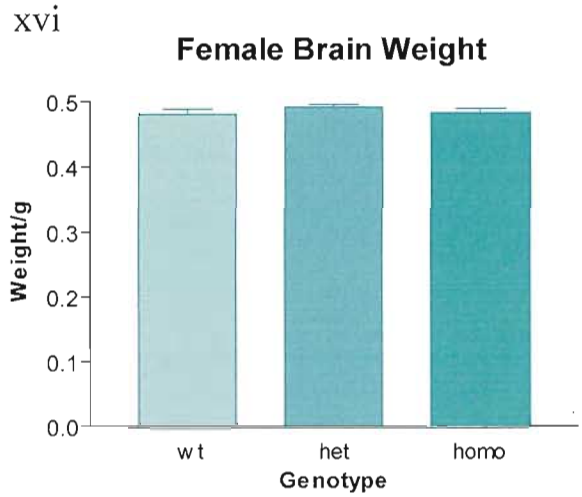
xiv



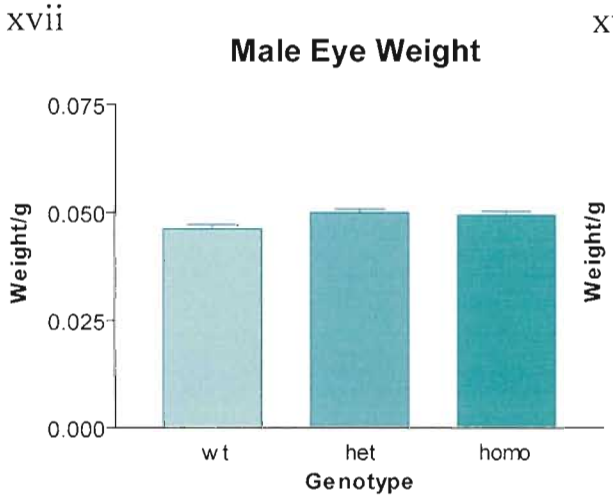
xv



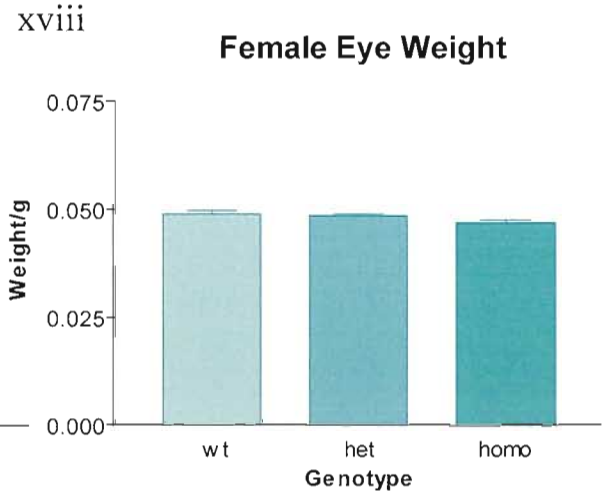
xvi



xvii

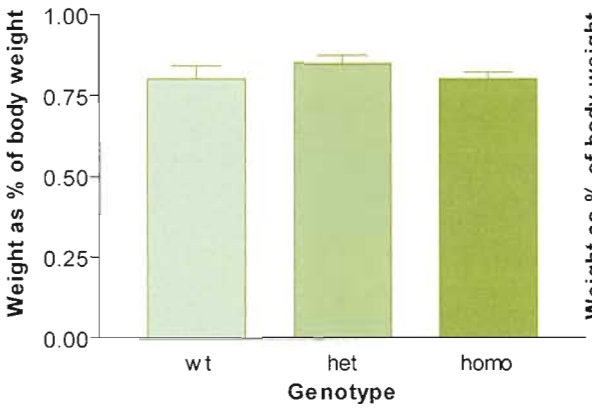


xviii

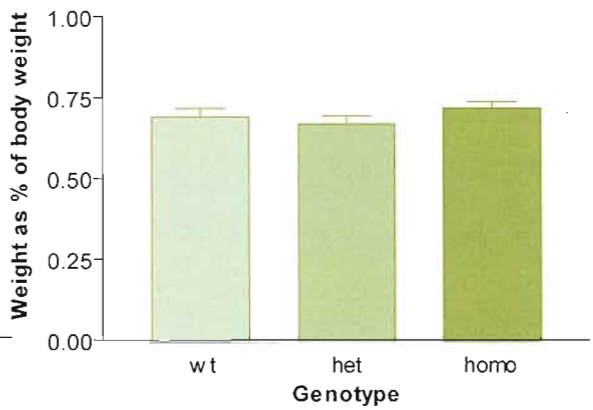


C

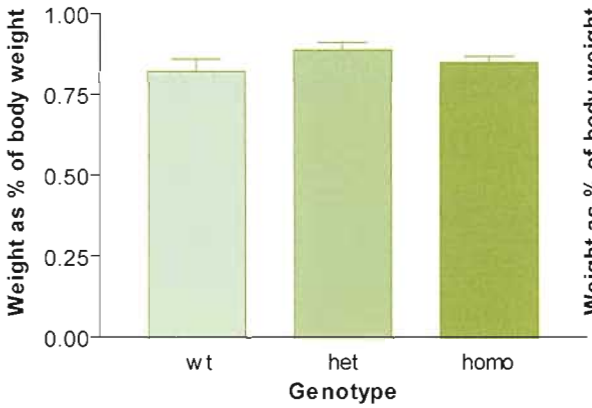
i



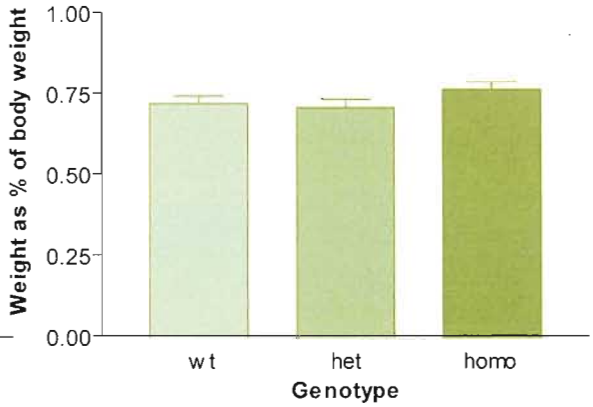
ii



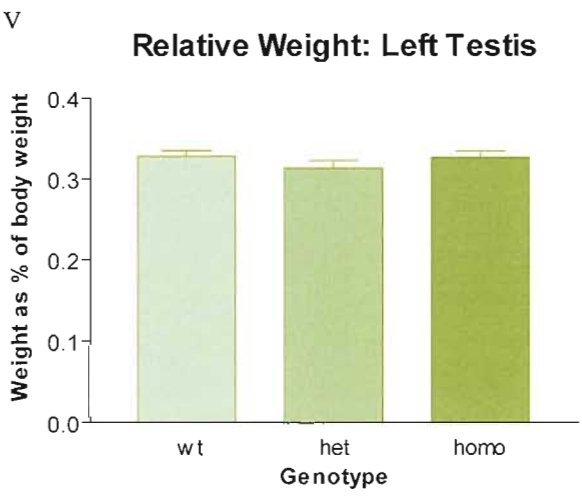
iii



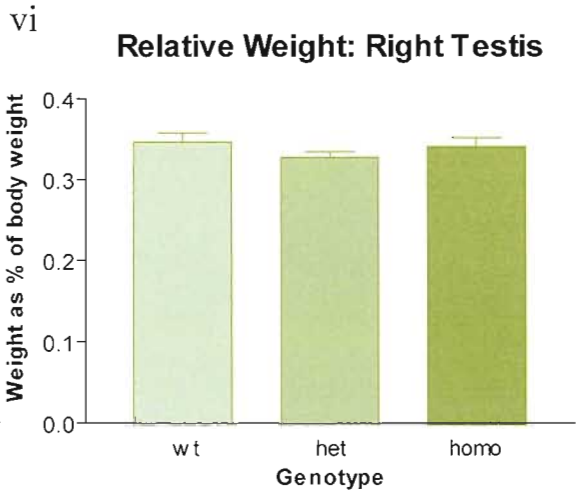
iv



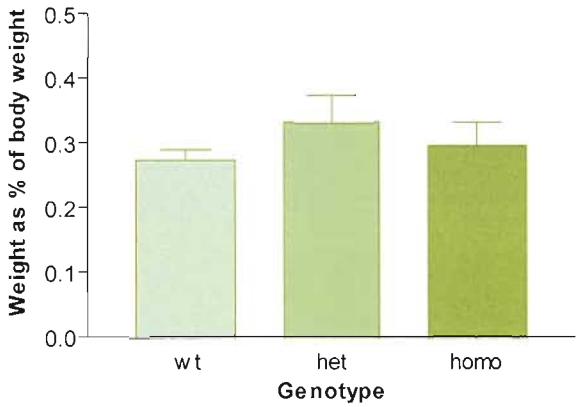
v



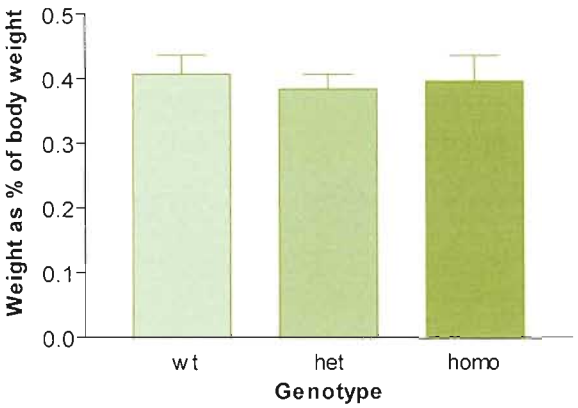
vi



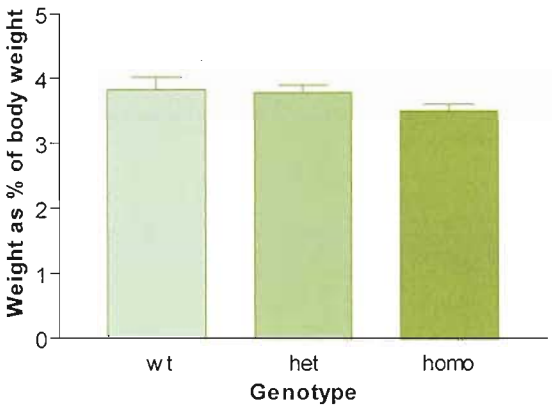
vii



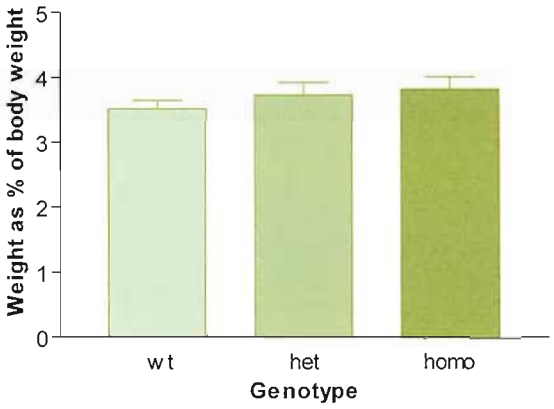
viii



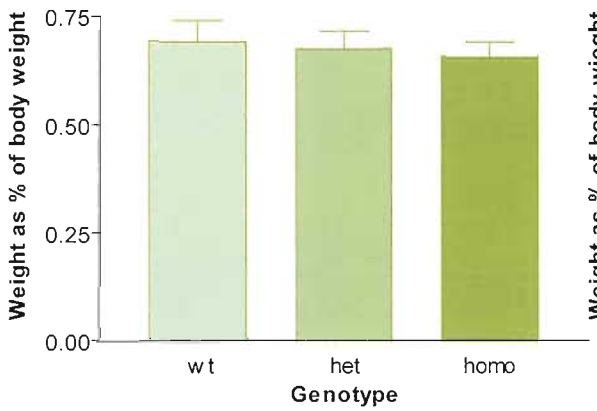
ix



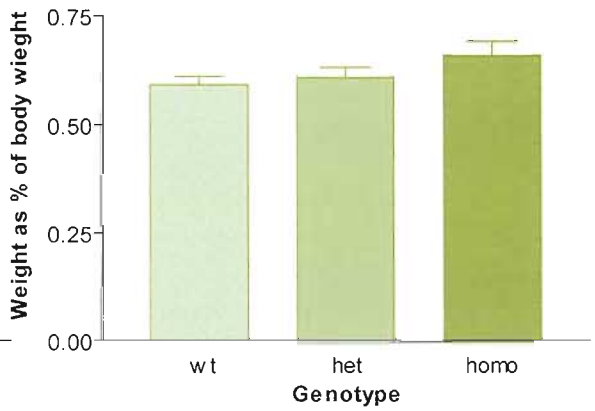
x



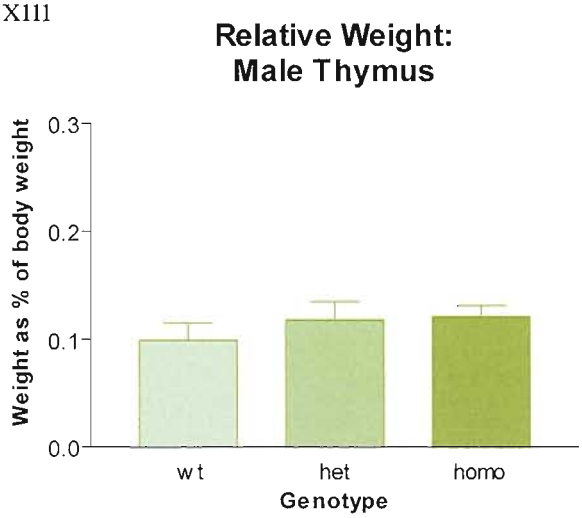
xi



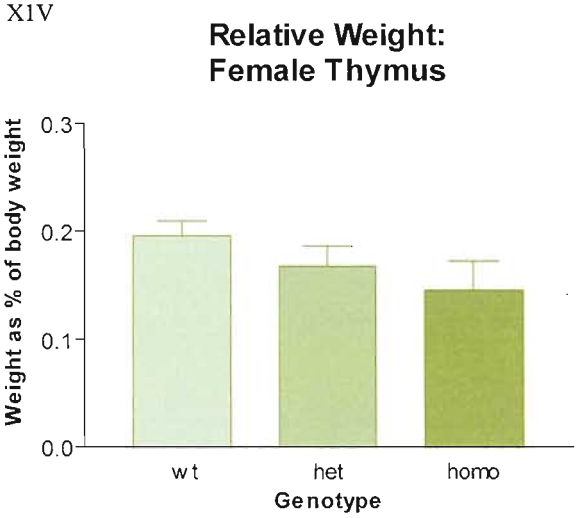
xii



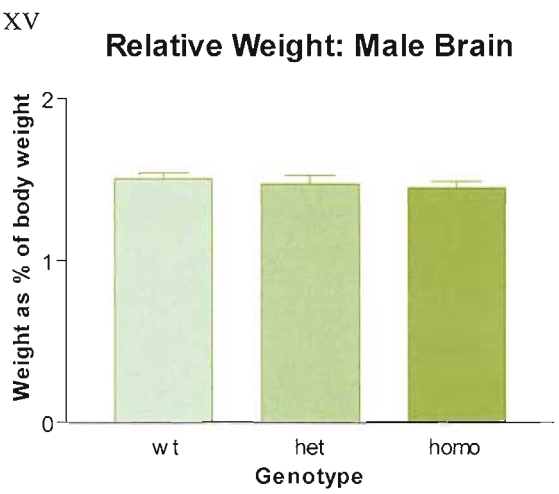
xiii



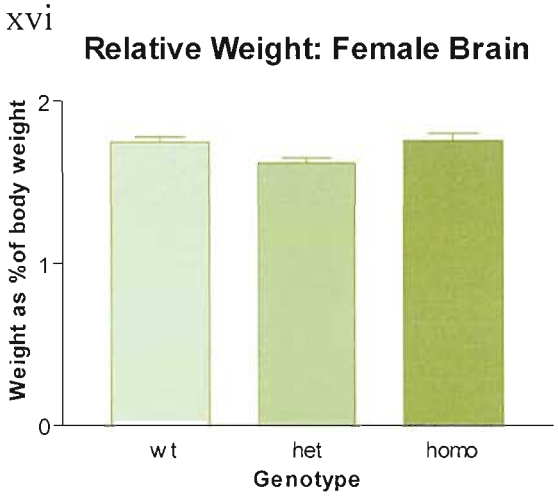
xiv



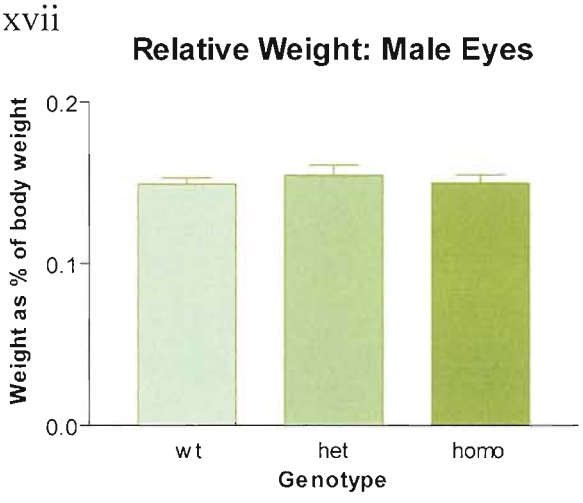
xv



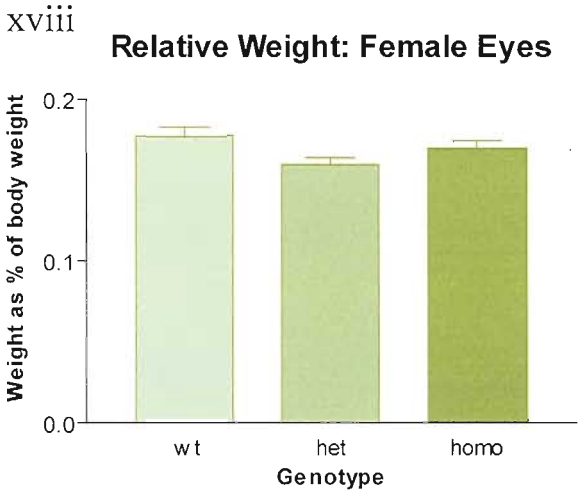
xvi



xvii



xviii



Liver, heart and thymus

Graphs showing liver weights can be seen in Figure 5.11B ix-x, and showing relative liver weights can be seen in Figure 5.11C ix-x. Graphs showing heart weights can be seen in Figure 5.11B xi-xii, and showing relative heart weights can be seen in Figure 5.11C xi-xii. Graphs showing thymus weights can be seen in Figure 5.11B xiii-xiv, and showing relative thymus weights can be seen in Figure 5.11C xii-xiv. There were no significant differences between the groups.

Brain

Graphs showing brain weights can be seen in Figure 5.11B xv-xvi, and showing relative brain weights can be seen in Figure 5.11C xv-xvi. There were no significant differences between the groups of males. There were also no significant differences between the groups of females with the actual weights. However, P=0.0247 for female relative brain weights. Tukey's post test shows that P<0.05 between wt and het, and also between het and homo. Het brains being smaller than both wt and homo.

Eyes

Graphs showing eye weights can be seen in Figure 5.11B xvii-xviii, and showing relative eye weights can be seen in Figure 5.11C xvii-xviii. P=0.0131 for male actual weights, whereas there was no significant differences between the groups of males with the relative weights. Conversely, there were no significant differences between the groups of females with the actual weights, whereas P=0.0395 for female relative weights. Tukey's post test shows for the males P<0.05 between wt and het, and between wt and homo; wt being smaller than both het and homo. Tukey's post test shows for the females P<0.05 between wt and het only; wt being larger than het.

- 1 year study

Body

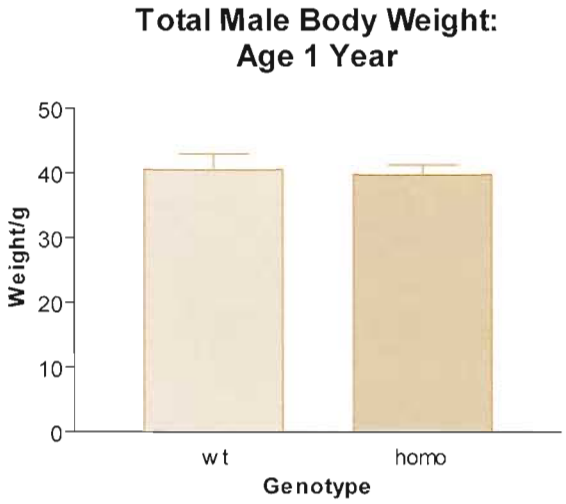
Firstly, total body weights of the males were weighed and analysed (Figure 5.12A). P=0.7956, therefore there was no significant difference between the body weights of the two groups.

Kidneys, spleen, heart and brain

Graphs showing kidney weights can be seen in Figure 5.12B i-ii, and showing relative kidney weights can be seen in Figure 5.12C i-ii. Graphs showing spleen weights can be seen in Figure 5.12B v, and showing relative spleen weights can

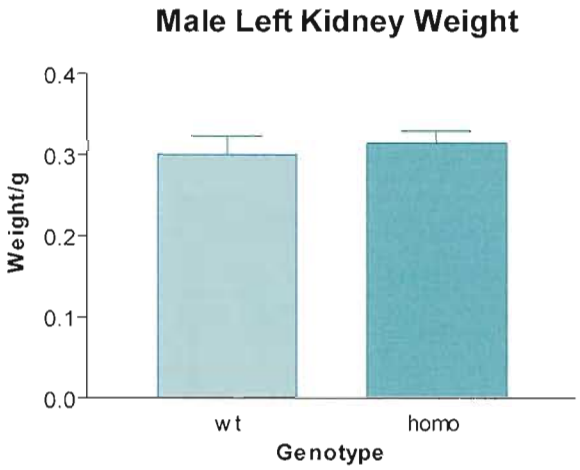
Figure 5.12. Graphs representing body and organ weights of 1 year old mice. (A) Graph showing body weights in grams. (B) Graphs showing organ weights in grams. (C) Graphs showing relative organ weights (organ weight as a percentage of total body weight). The sample populations are 129 wildtype (wt) and homozygote (homo) 1 year old male mice born from a Travis heterozygous intercross. (B) i-x: Graphs showing organ weights in grams. i, left kidney; ii, right kidney; iii, left testis; iv, right testis; v, spleen; vi, liver; vii, heart; viii, thymus; ix, brain; x, eyes. (C) i-x: Graphs showing relative organ weights (graphs in the same order as in B).

A

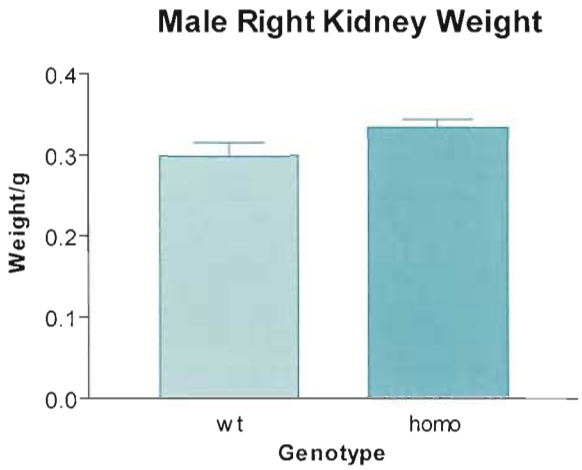

Total Male Body Weight:
Age 1 Year
Weight/g
50
40
30
20
10
0
wt
homo
Genotype

B

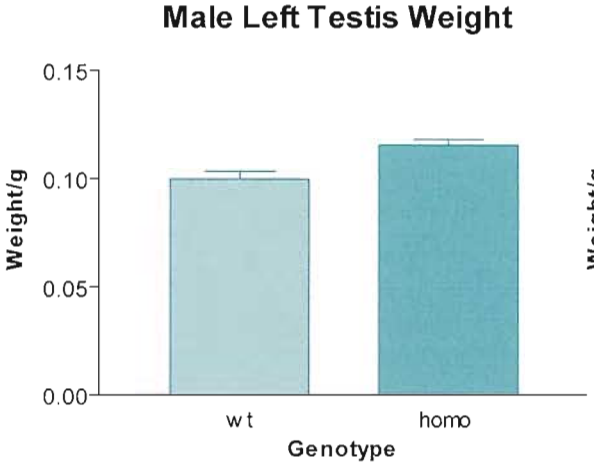
i


Male Left Kidney Weight
Weight/g
0.4
0.3
0.2
0.1
0.0
wt
homo
Genotype

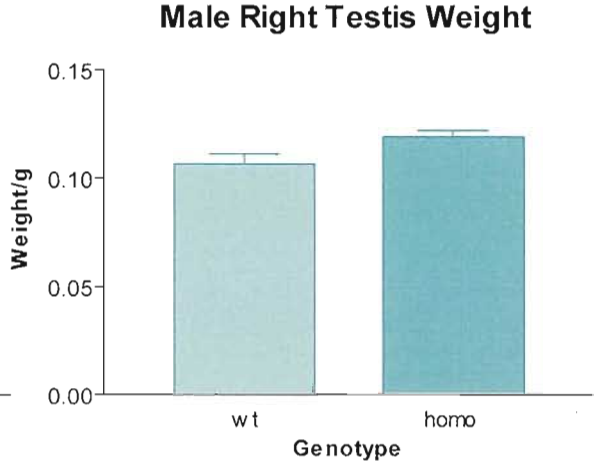
ii


Male Right Kidney Weight
Weight/g
0.4
0.3
0.2
0.1
0.0
wt
homo
Genotype

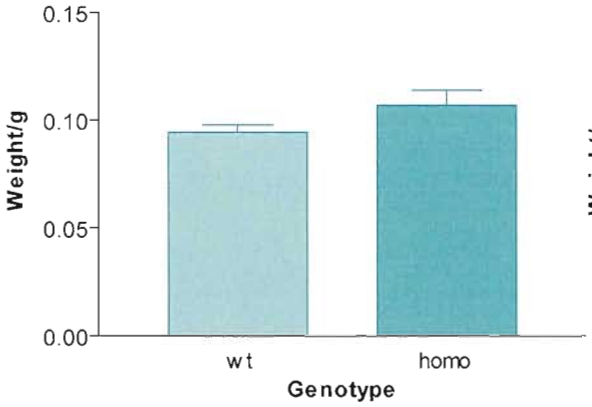
iii


Male Left Testis Weight
Weight/g
0.15
0.10
0.05
0.00
wt
homo
Genotype

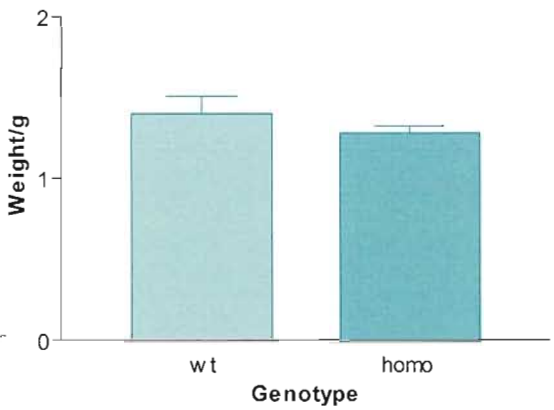
iv


Male Right Testis Weight
Weight/g
0.15
0.10
0.05
0.00
wt
homo
Genotype

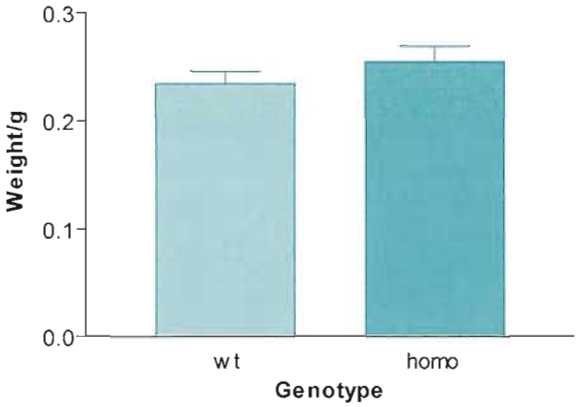
v



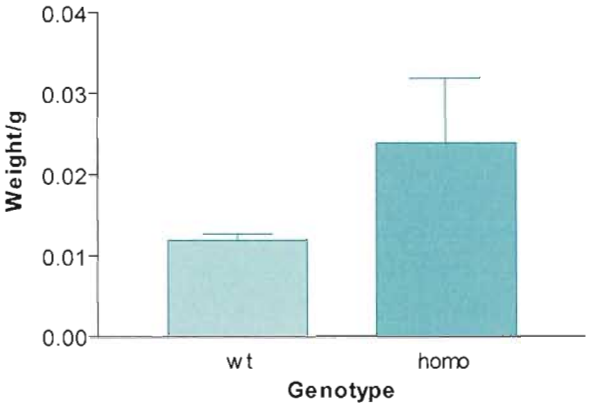
vi



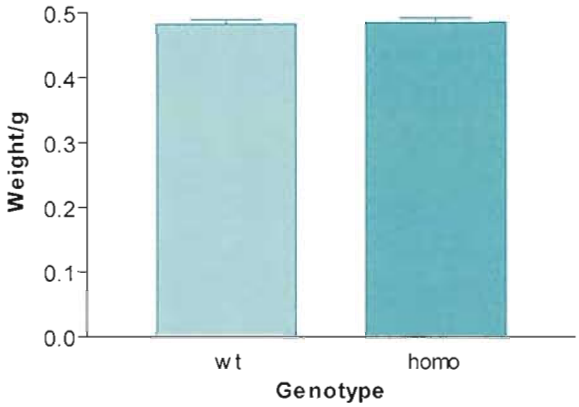
vii



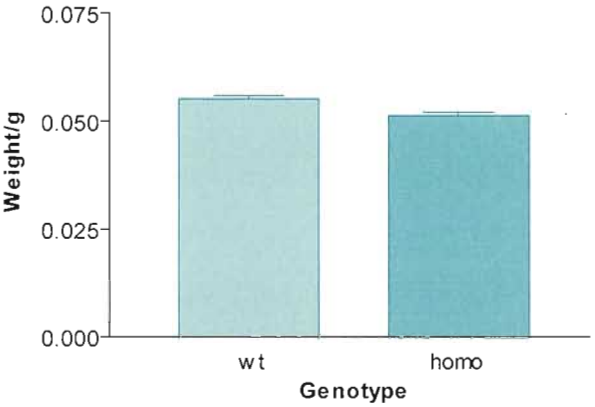
viii



ix

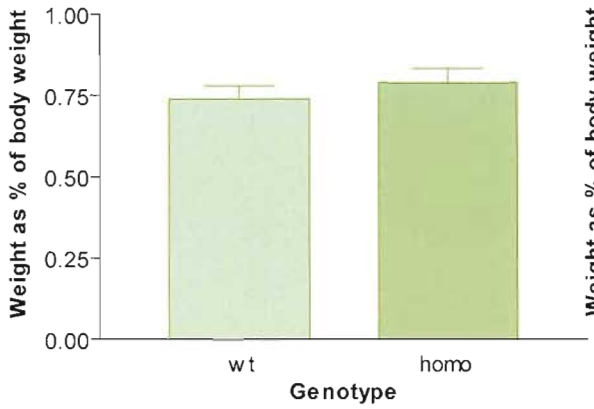


x

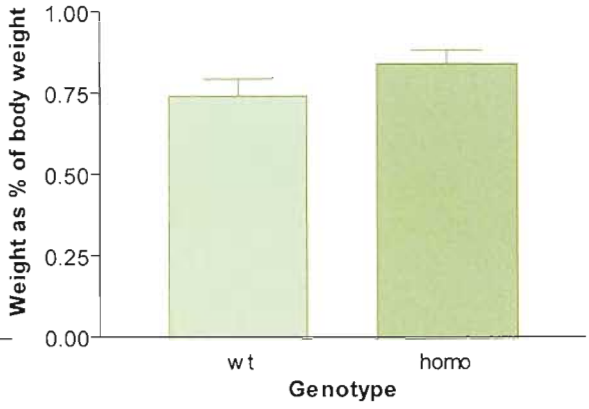


C

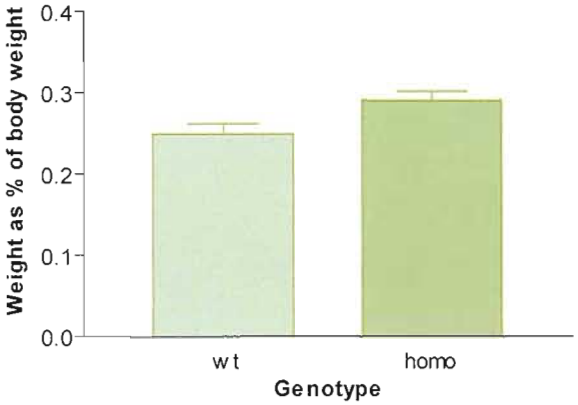
i



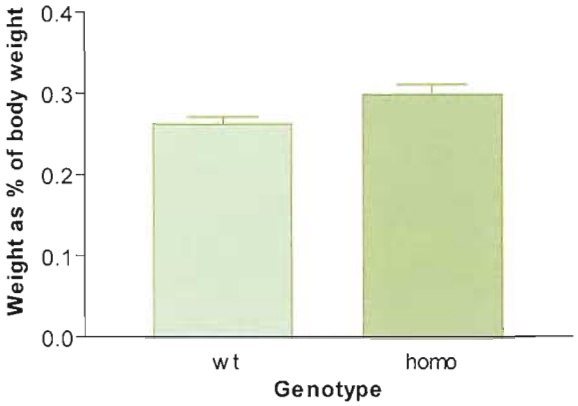
ii



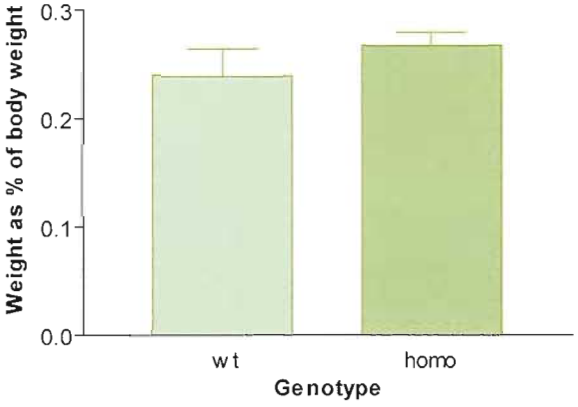
iii



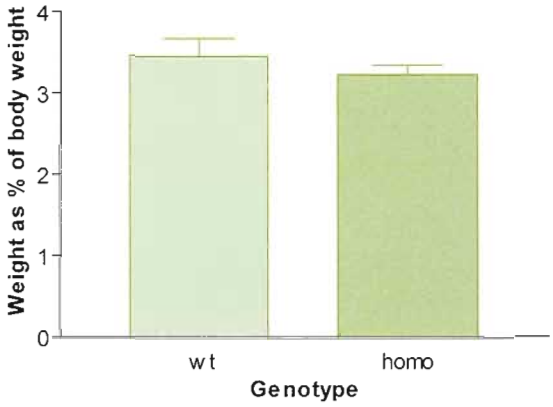
iv



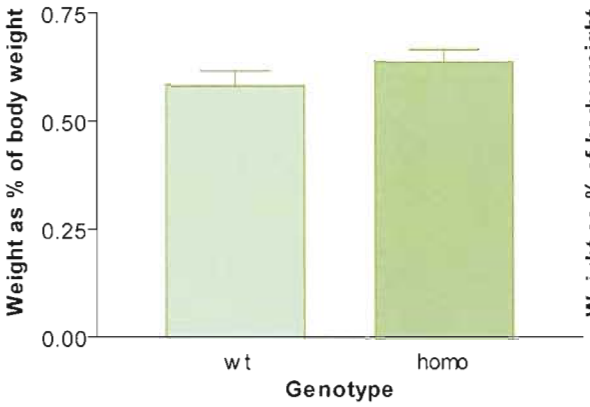
v



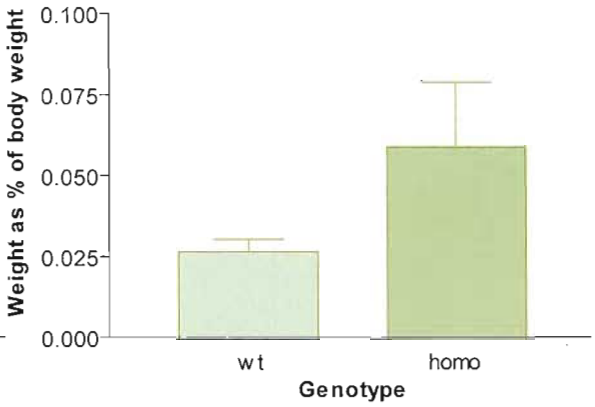
vi



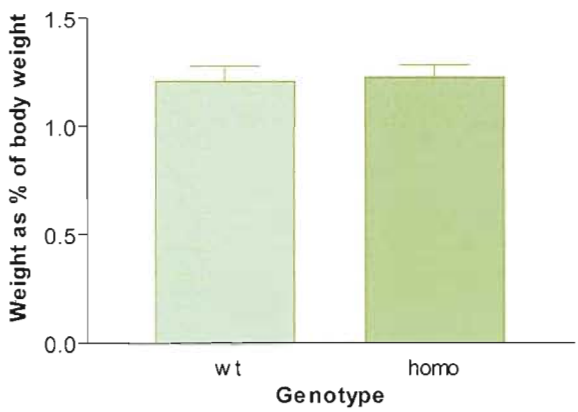
vii


Relative Weight: Male Heart
Weight as % of body weight
0.75
0.50
0.25
0.00
wt
homo
Genotype

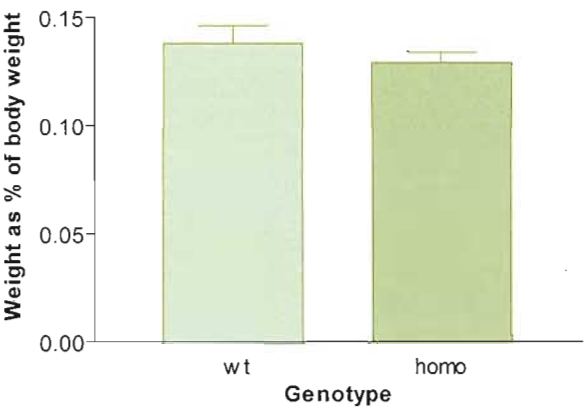
viii


Relative Weight: Male Thymus
Weight as % of body weight
0.100
0.075
0.050
0.025
0.000
wt
homo
Genotype

ix


Relative Weight: Male Brain
Weight as % of body weight
1.5
1.0
0.5
0.0
wt
homo
Genotype

x


Relative Weights: Male Eyes
Weight as % of body weight
0.15
0.10
0.05
0.00
wt
homo
Genotype

be seen in Figure 5.12C v. Graphs showing heart weights can be seen in Figure 5.12B vii, and showing relative heart weights can be seen in Figure 5.12C vii. Graphs showing brain weights can be seen in Figure 5.12B ix, and showing relative brain weights can be seen in Figure 5.12C ix. There were no significant differences between the two groups.

Testes

Graphs showing testes weights can be seen in Figure 5.12B iii-iv, and showing relative testes weights can be seen in Figure 5.12C iii-iv. Both left and right testes were significantly different between the two groups, for both actual (P=0.0051, P=0.0461) and relative weights (P=0.0306, P=0.0263). Homo being larger than wt.

Liver

Graphs showing liver weights can be seen in Figure 5.12B vi, and showing relative liver weights can be seen in Figure 5.12C vi. There was no significant difference between the relative liver weights of the two groups. The F test to compare variances of the actual liver weights had a P value <0.05. Since the t-test assumes equal variance between groups the data couldn't be used with a standard t-test. A standard t-test calculates the degrees of freedom to be the population size (n) – 2. A Welch's correction was carried out which calculates the degrees of freedom in a more complex way than a standard t-test. There was no significant difference between the actual liver weights of the two groups with a Welch's t-test modification.

Thymus

Graphs showing thymus weights can be seen in Figure 5.12B viii, and showing relative thymus weights can be seen in Figure 5.12C viii. The F test to compare variances of the actual and relative thymus weights had a P value <0.05. Following a Welch's t-test modification there was no significant difference between the thymus weights of the two groups.

Eyes

Graphs showing eye weights can be seen in Figure 5.12B x, and showing relative eye weights can be seen in Figure 5.12C x. There was no significant difference between the relative eye weights of the two groups. P=0.0039 for the actual eye weights; therefore there was a significant difference between the two groups. Homo being smaller than wt.

5.4.12.3 Histological analysis

- 7-8 month study

Six samples of kidney, testis, liver, heart, brain and eye of the organs dissected from the 7-8 month old cohort were sectioned and histologically examined. These six samples were from a male and female wildtype; a male and female heterozygote; and a male and female homozygote. With the testes, two male samples from each genotype were chosen. The kidney, testis, spleen, thymus, brain and eye were chosen to be examined as they are all sites of *Wt1* expression (see section 1.4.7); therefore any affect on *Wt1* due to the aberrant expression of *Wt1-AS* in these Travis mice may have had a physical affect on each of these organs. The liver and heart were chosen to be examined as *Wt1* is expressed in the pleural and pericardial space surrounding these respective organs, so again, any affect on *Wt1* due to the aberrant expression of *Wt1-AS* in these Travis mice may have had a physical affect on these organs. All sections apart from kidneys were H&E stained; kidneys were PAS stained.

Kidneys

Kidneys were stained with PAS to visualise glycoproteins in order to distinguish between the proximal and distal convoluted tubules; proximal convoluted tubules have pink staining within the brush borders whereas distal convoluted tubules have no staining. All structures of the kidney were present in each of the samples and their distribution and density within the cortex and medulla appeared normal (Figure 5.13A).

Testes

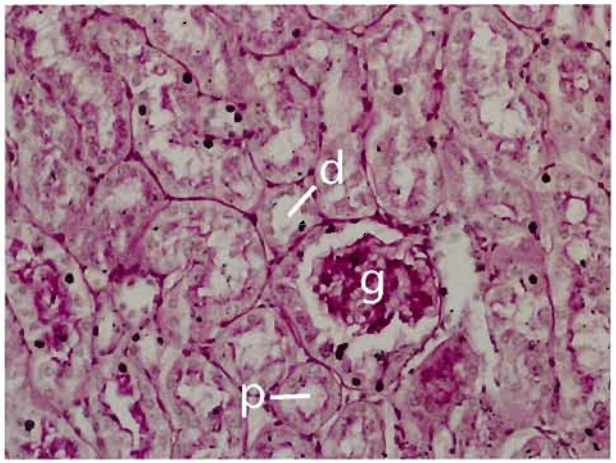
The size and distribution of seminiferous tubules and Leydig cells within all samples appeared to be the same (Figure 5.13B). Under high magnification (400x), the different stages of spermatogenesis could be seen, although it was difficult to distinguish the Sertoli cells from immature spermatogonia in any of the samples due to the resolution power of the microscope.

Two heterozygous mice exhibited abnormal testes and although excluded from the weight study, histology was carried out (Figure 5.13G). The first testis (Figure 5.13Gi) was abnormally small in size, although the other testis of the

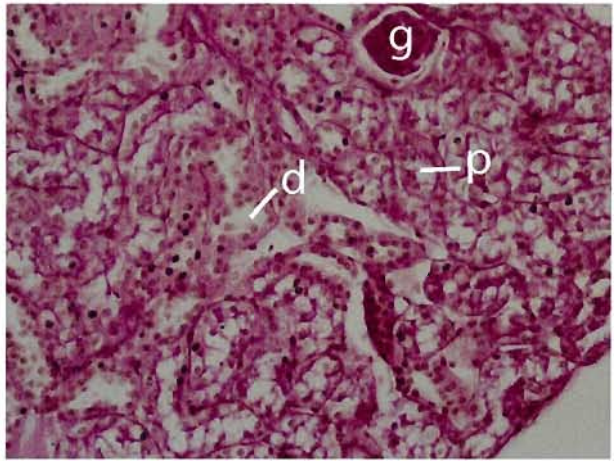
Figure 5.13. (A-G) Histological organ sections of 7-8 month old mice born from a Travis heterozygous intercross. (A) PAS stained kidney sections (200x). (B) H&E stained testis sections. i, ii, iii (100x); vi, v, vi (400x). (C) H&E stained liver sections (200x). (D) H&E stained heart sections (200x). (E) H&E stained brain sections; choroid plexus in the region of the fourth ventricle (200x). (F) H&E stained eye sections (200x). (G) H&E stained abnormal testis sections (100x) and abnormal liver sections (200x). (A, C, D, E, F) i, male wildtype; ii, male heterozygote; iii, male homozygote; iv, female wildtype; v, female heterozygote; vi, female homozygote. (B) i, iv, male wildtype; ii, v, male heterozygote; iii, vi, male homozygote. (G) i, ii, testis sections; iii, iv, liver sections. d, distal convoluted tubule; p, proximal convoluted tubule; c, comma-shaped body; s, s-shaped body; g, glomerulus; a, interlobular artery; sf, seminiferous tubule; l, leydig cells; m, heart muscle; c, choroid plexus; v, fourth ventricle; r, roof of fourth ventricle; gcl, ganglion cell layer; ipl, inner plexiform layer; inl, inner nuclear layer; opl, outer plexiform layer; onl, outer nuclear layer; rc, layer of rods and cones; e, pigment epithelium; ch, choroid.

A

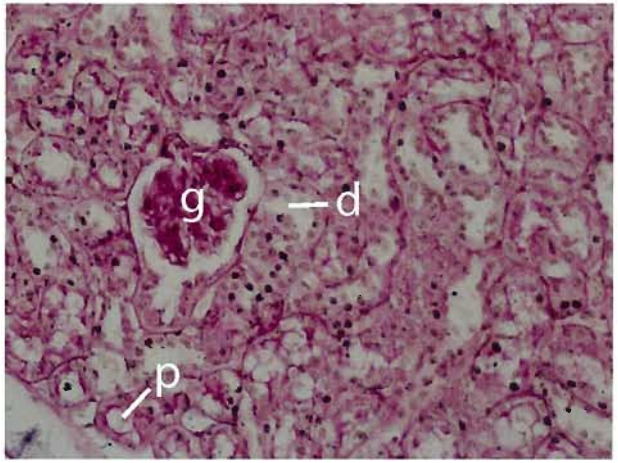
i



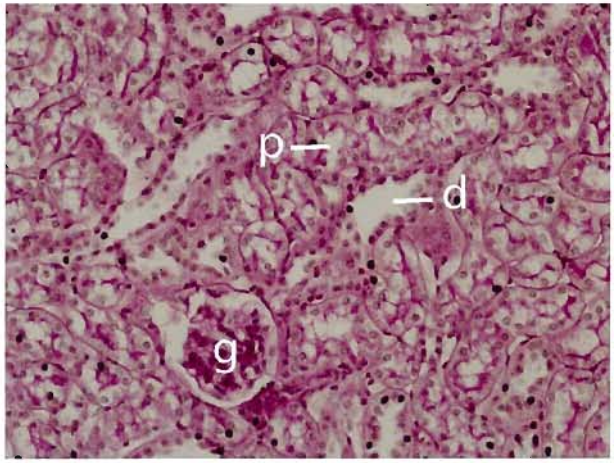
ii



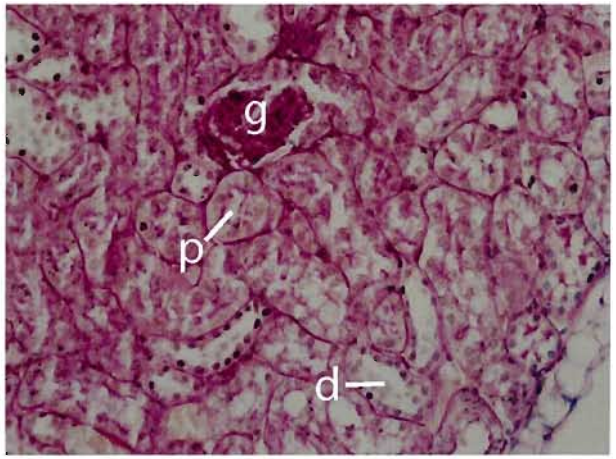
iii



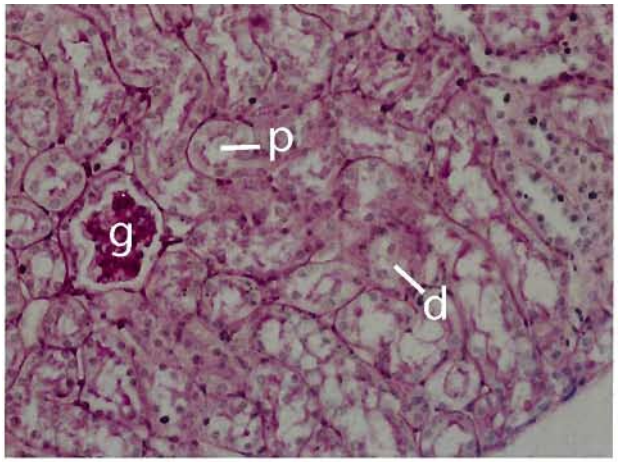
iv



v

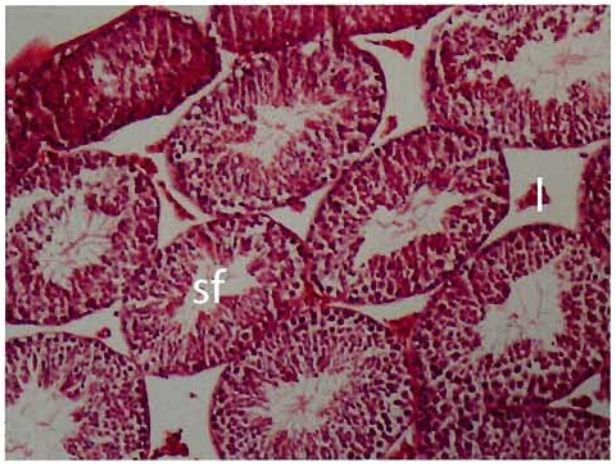


vi

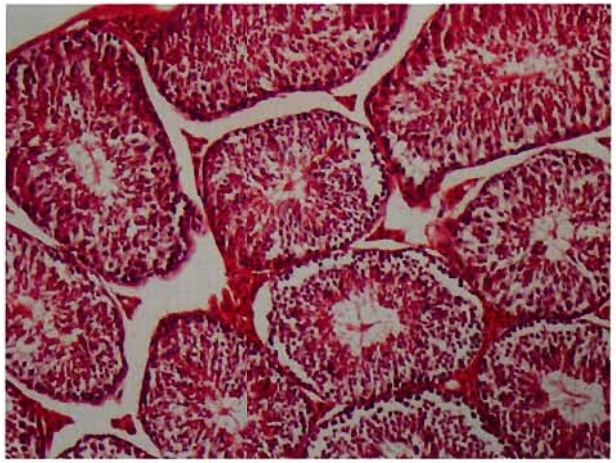


B

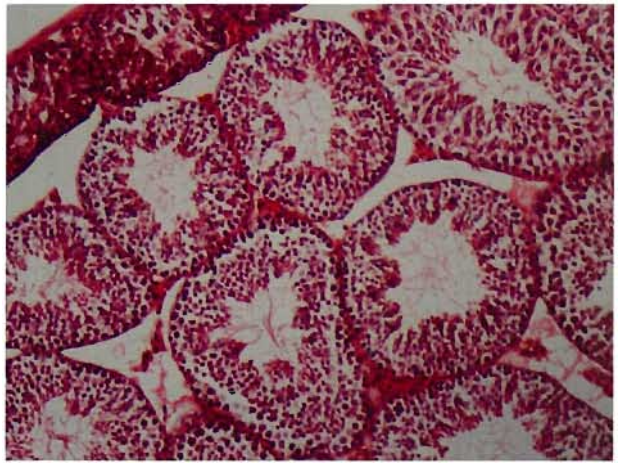
i



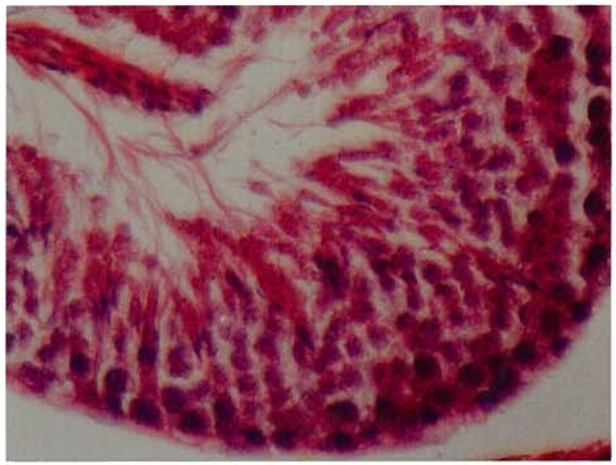
ii



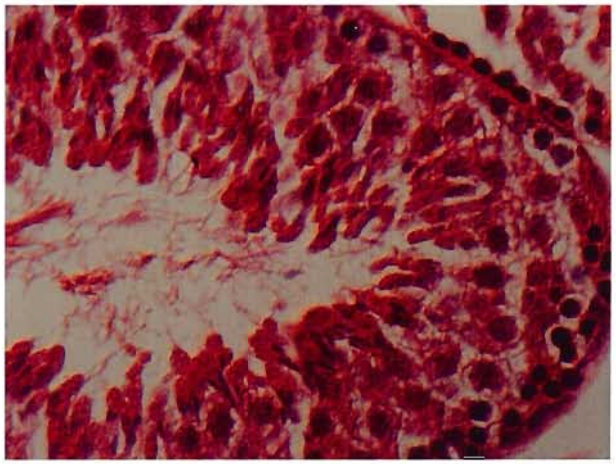
iii



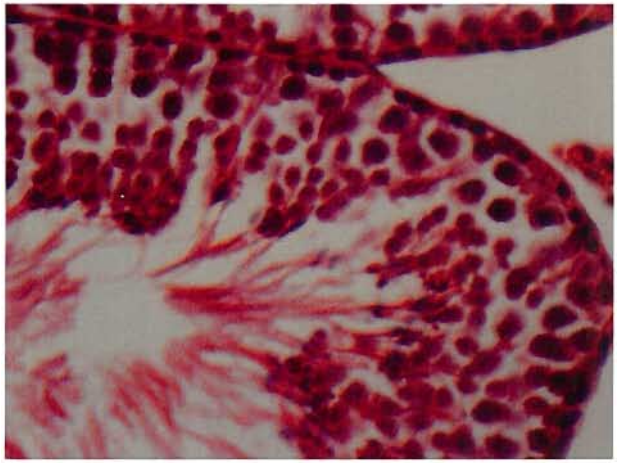
iv



v

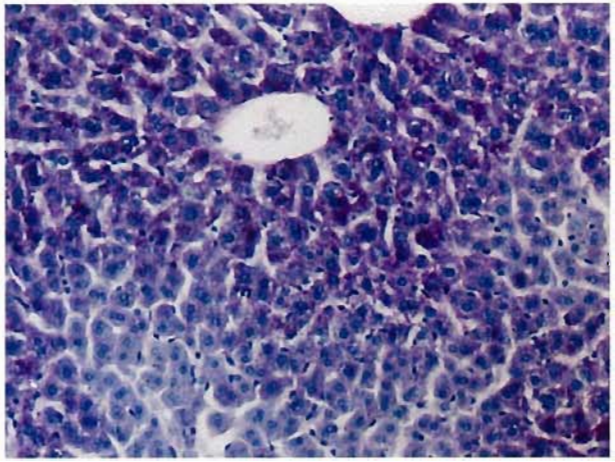


vi

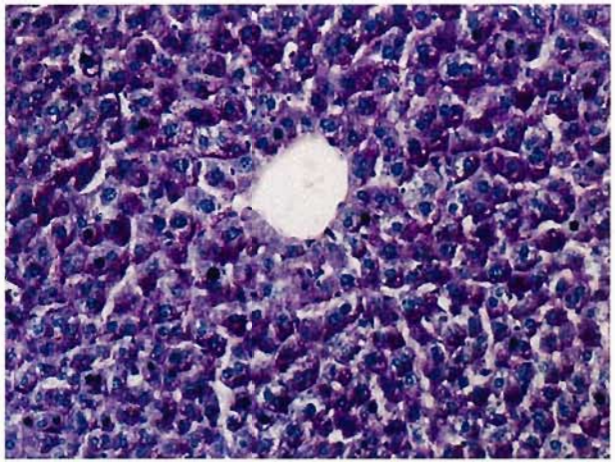


C

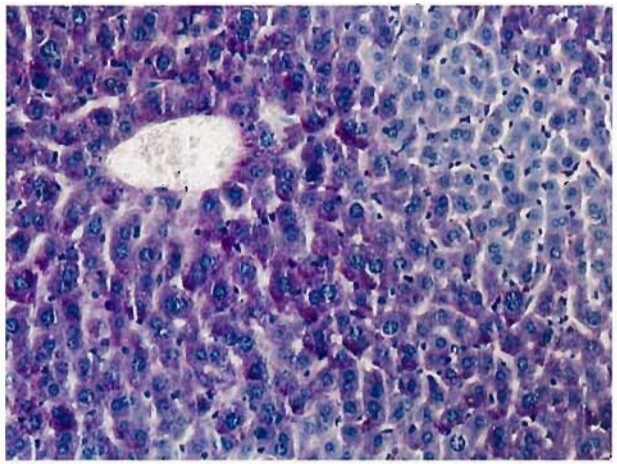
i



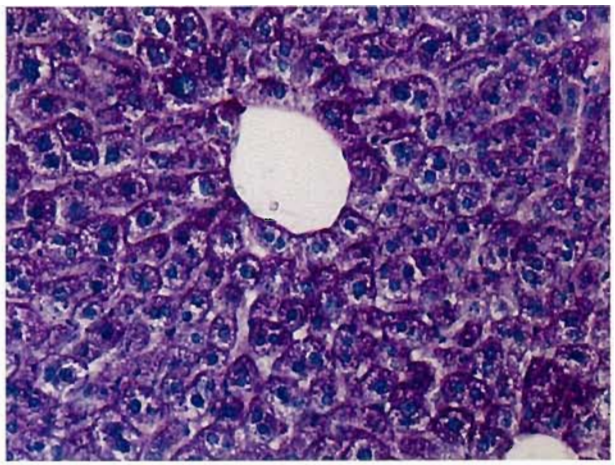
ii



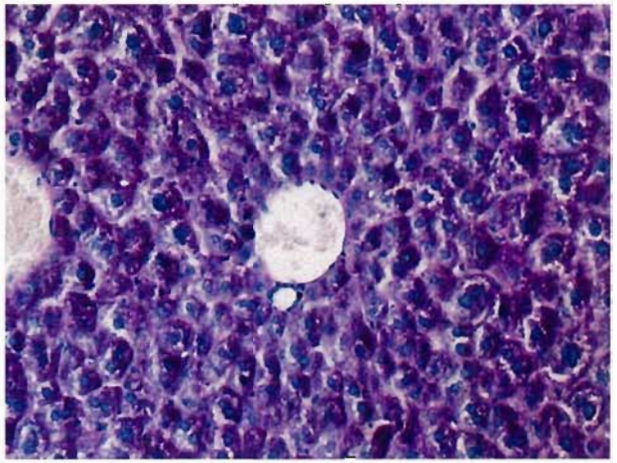
iii



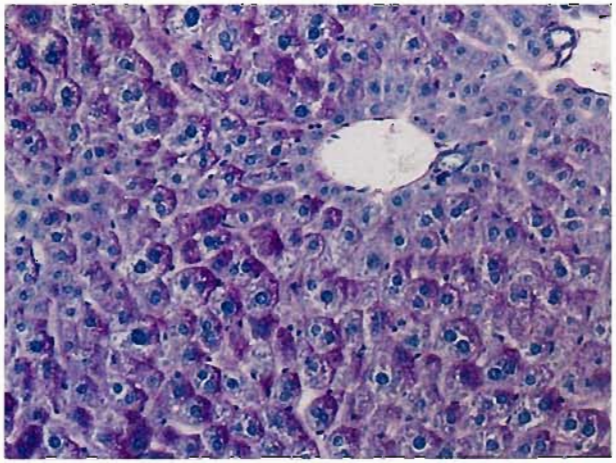
iv



v

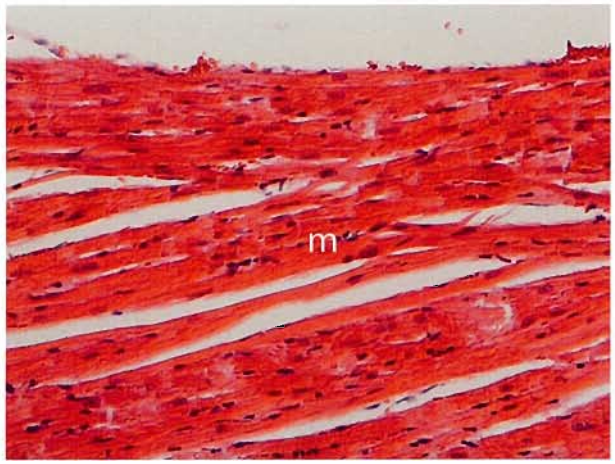


vi

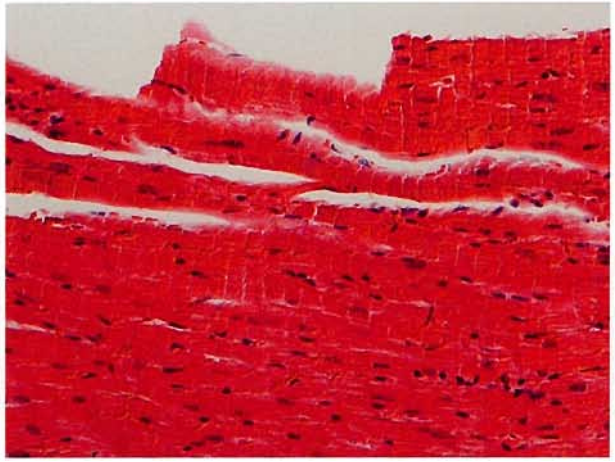


D

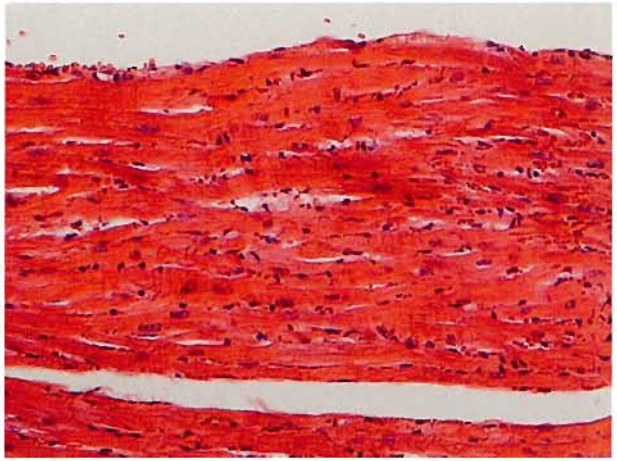
i



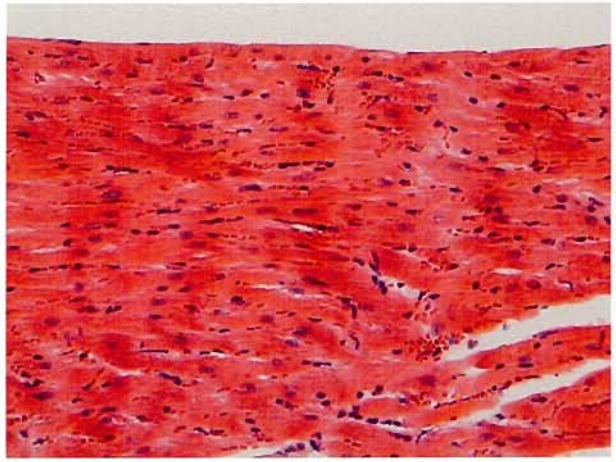
ii



iii

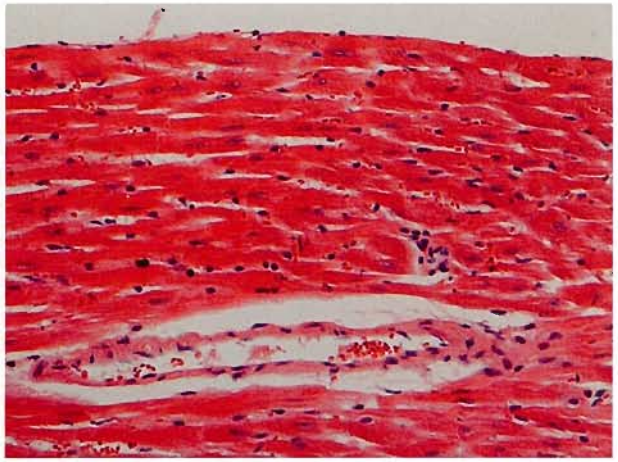


iv



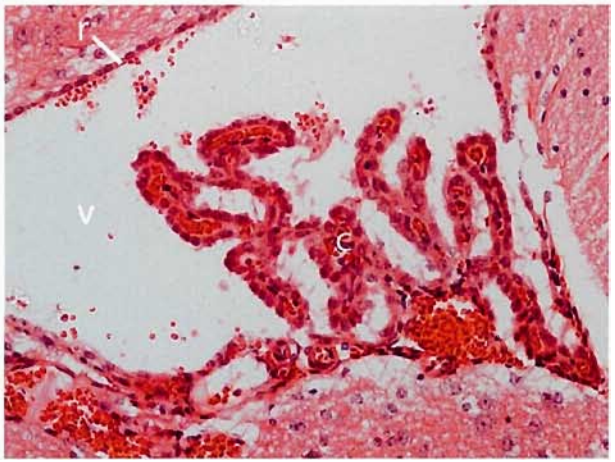
v

vi

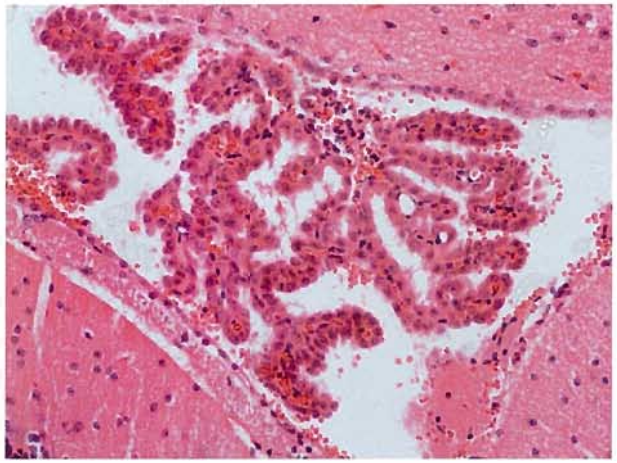


E

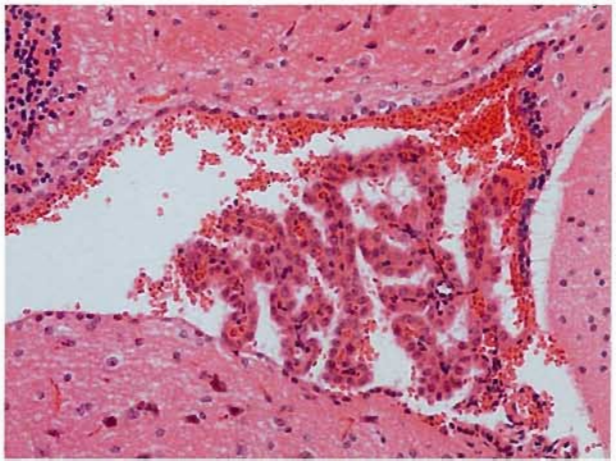
i



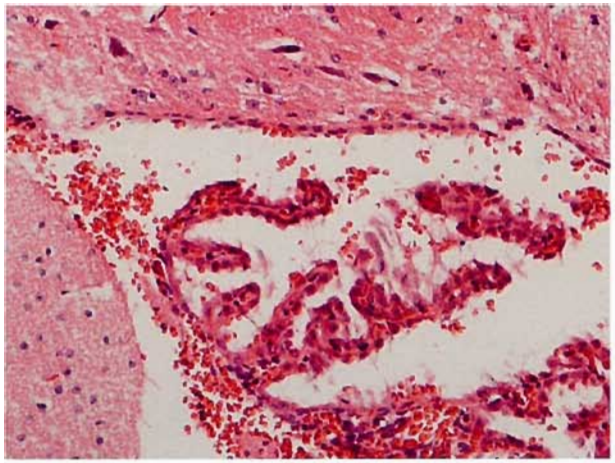
ii



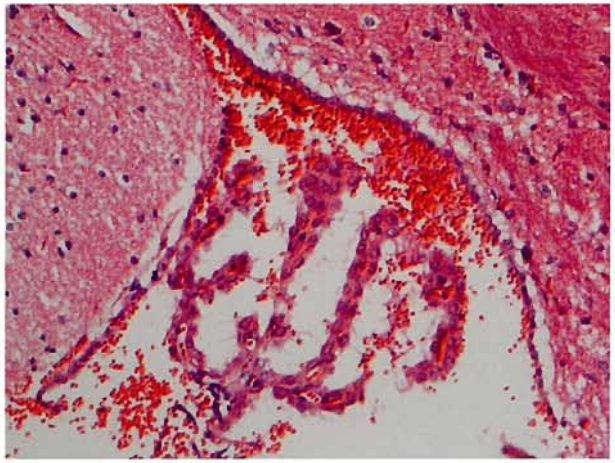
iii



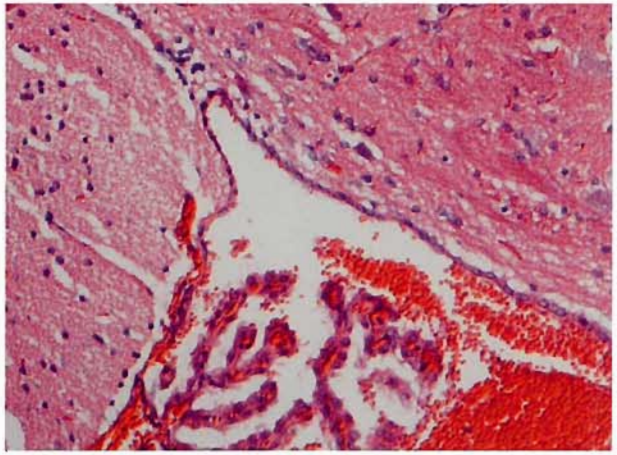
iv



v

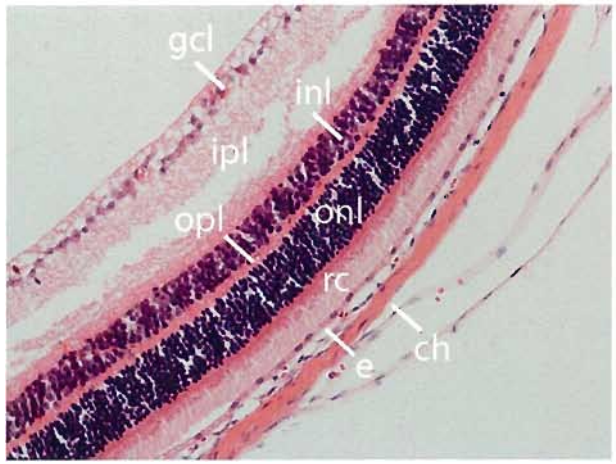


vi

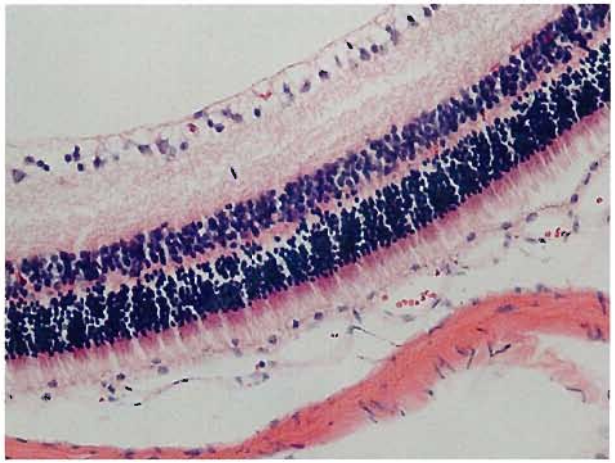


F

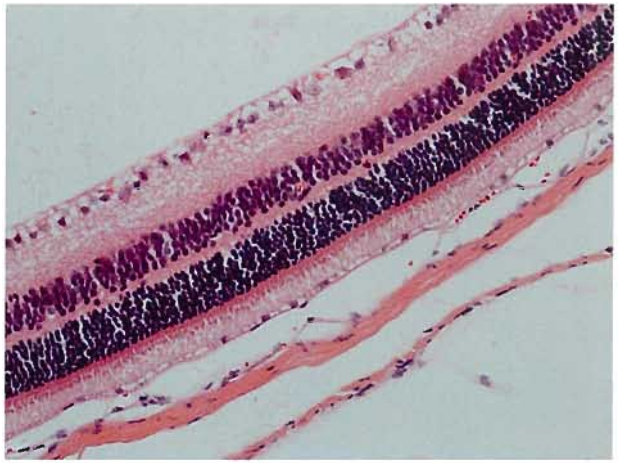
i



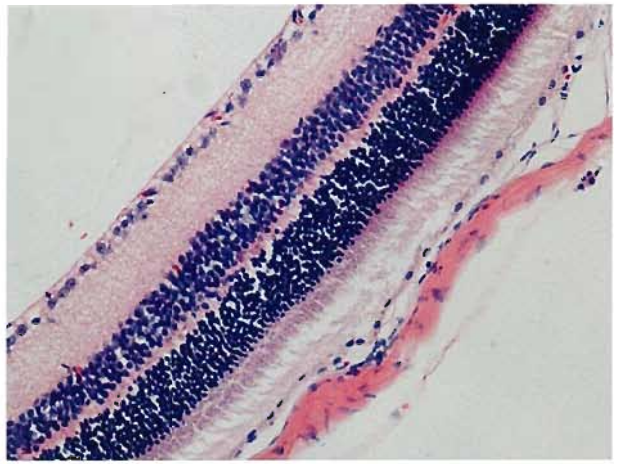
ii



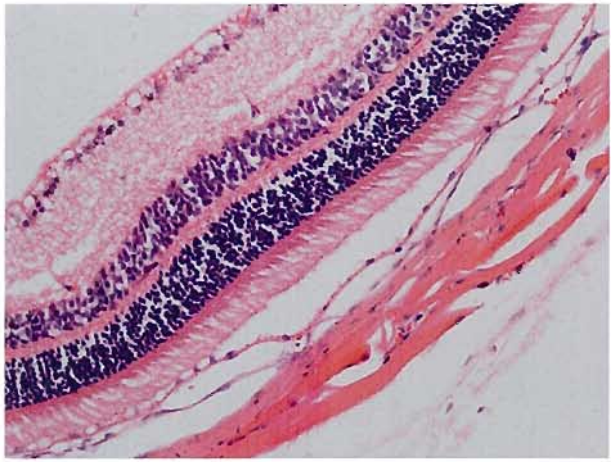
iii



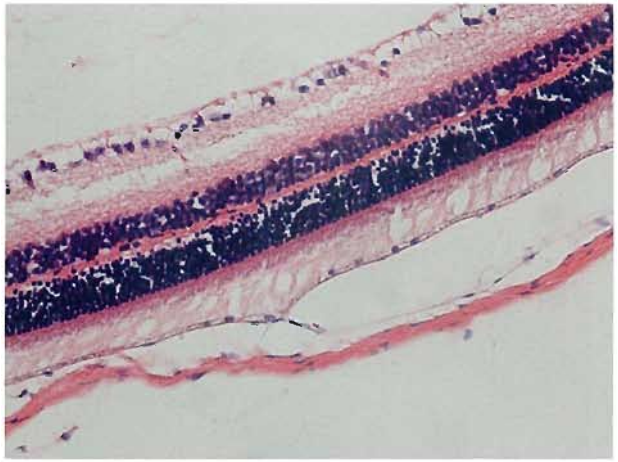
iv



v

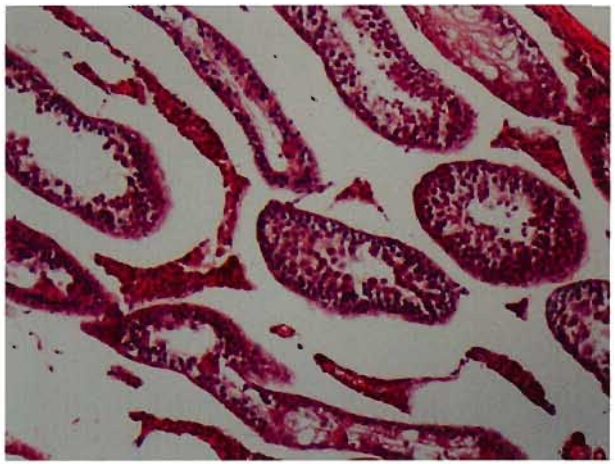


vi

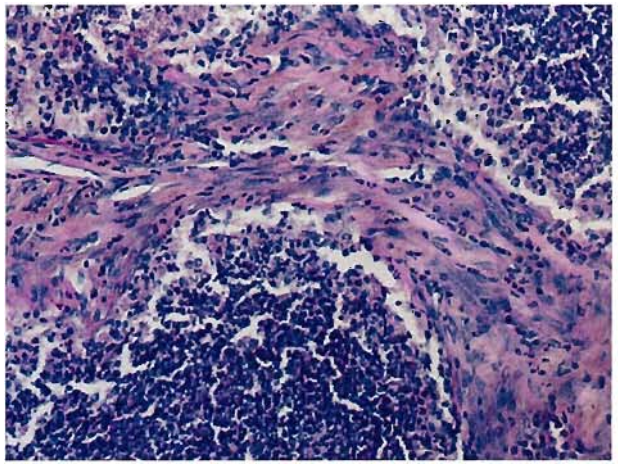


G

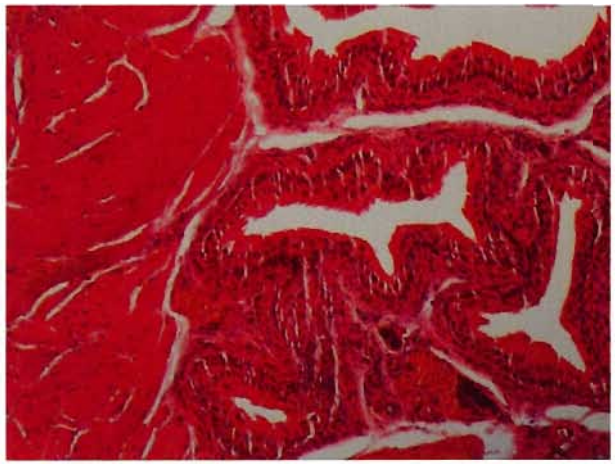
i



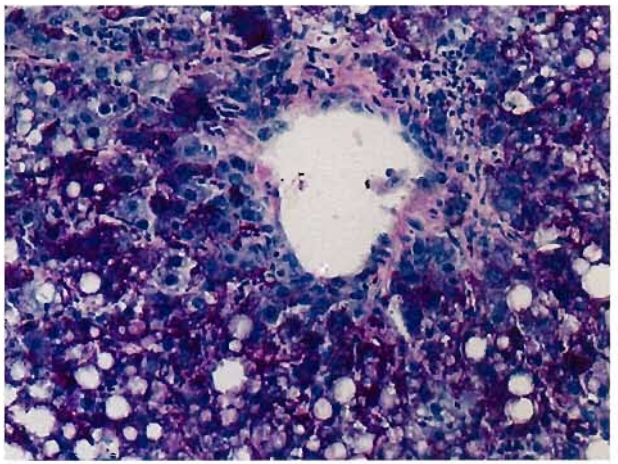
iii



ii



iv



same animal appeared normal. Leydig cells and seminiferous tubules with what looked like mature spermatogonia were present. The empty space between the seminiferous tubules however was much larger than normal. The second testis (Figure 5.13Gii) was abnormally shaped and small in size, although again, the other testis of the same animal appeared normal. Histologically, this tissue did not resemble testis.

Liver

The distinctive radial pattern of hepatocytes surrounding a central vein can be seen in all samples (Figure 5.13C). There does not seem to be any disorganisation of cells or differences between cell size between the samples.

One wildtype female mouse and one male heterozygous mouse were found to have grossly abnormal livers and although excluded from the weight study, histology was carried out (Figure 5.13G). Histologically, both livers exhibited abnormal cell morphology characteristic of tumours.

Heart

Sections through whole hearts were initially visualised at 12.5x magnification in order to try and observe each of the chambers to look for obvious abnormalities of structure or size (data not shown). Due to variations in the orientation of the hearts during embedding, a comprehensive comparison between the samples was not possible as sections from each sample showed the chambers from a different plane of view. However, no gross abnormalities were observed. Sections were also visualized at a higher magnification in order to compare the muscle layers of the ventricles of the heart. There were no obvious differences between the muscle layers between the samples (Figure 5.13D).

Brain

Wt1 expression in the brain is known to be restricted to the roof of the fourth ventricle and an area within the fourth ventricle, the area postrema (Armstrong et al., 1993; Sharma et al., 1992). Serial brain sections were stained and visualised until the fourth ventricle region was found. The cells lining the roof of the fourth ventricle displayed an organised morphology and the fluid filled region of the ventricular space appeared to be of a comparable size for each sample (Figure 5.13E).

Eyes

Wt1 murine expression in the eye is known to occur in the retina and developing lens vesicle at E12 and the presumptive retinal ganglion layer between E15-PD1 (Wagner et al., 2002b). Additionally, on an MF1:C57BL/6 F1 hybrid background, *Wt1* homozygous null mutants have thinner than normal retinas, apoptosis of retinal ganglion cells and disruption of optic nerve fibre growth (Wagner et al., 2002b).

The orientation of the eye was an important consideration when embedding the samples as the photoreceptor cells of the retina change in number depending on where in the retina layer they are located. Cones cells for example are more numerous within the central retina in the region containing the fovea centralis and surrounding macula. Rod cells on the other hand are more numerous in the peripheral retina and completely absent from the fovea centralis. Difficulties with orientation arose during the embedding of the eyes due to the strain background of the mice used (129 strain). The eyes within this strain were white/translucent and it was not always possible to accurately ascertain the orientation of eye during embedding. When the orientation could be established due to the presence of a small amount of remaining muscle attachment, the eyes were embedded with the lens facing upwards. To circumvent the problem of comparing different regions of the retina due to the random orientation of the eyes during embedding, serial sections were stained and only planes of view containing both the lens and retina were compared. The retina layer adjacent to the lens was observed for each sample as the retina layer closest to the lens would have been within the peripheral region; the peripheral retina being rod-dominated and the central retina being cone-dominated.

The overall width of the retina layer in the homozygous samples appeared thinner than that of the wildtype and heterozygous samples (Figure 5.13F), although it was not obvious that any one layer of the retina could account for this observation.

5.4.13 Travis mice exhibiting morbid features

As well as the cohort of animals that have been examined, there were several Travis mice that became ill and had to be culled or were found dead in their cage. A summary of these mice can be found in Table 5.6. Six of these mice (five heterozygotes, one wildtype) had abnormal growths, some of which were confirmed histologically as tumours (Figure 5.13G, other data not shown). When DNA could be extracted from these growths, they were genotyped to check for a conversion from a heterozygous genotype to a homozygous genotype which may explain the appearance of tumours. Two of these growths were genotyped, but no conversion from a heterozygous genotype to a homozygous genotype was found. Although it seems striking that no homozygotes appeared with abnormal growths, this may be explained by the breeding strategy whereby twice as many heterozygotes compared to wildtypes or homozygotes were born due to heterozygous intercrosses. Two mice that displayed loss of motor coordination were both part of the same litter and it was therefore possible that a spontaneous mutation within one of the parents could have been to blame, however, four other homozygous mice born by the same parents displayed no such phenotype.

5.5 Discussion

5.5.1 Choice of the MCRII region to mutate within the mouse genome

Two proposals for knocking out the *Wt1-AS* transcript were considered. The first being to knock out the promoter region and the second being to knock out the MCRII region whilst also truncating the *Wt1-AS* transcript. The second option was chosen for several reasons. The MCRII region is one of two highly conserved non-coding DNA elements between mouse and man, and was chosen over MCRI as it is closer to 5' end of the transcript. It has also been proposed to act as a regulator of *Wt1* expression (see section 1.4.10). This targeting strategy allows two questions to be answered: what is the effect of a truncated *Wt1* antisense transcript from the MCRII region onwards and what is the effect of a MCRII genomic deletion.

Had the promoter region of *Wt1-AS* been found, the targeting of this gene would have been more straight forward, as knocking out the promoter region would

have prevented any transcription of this gene. If the promoter region been narrowed down, our collaborators (CLIC Sargent Research Unit) could have carried out luciferase reporter assays by transient transfection with regions of murine DNA sequence in the antisense orientation, the technique used by Malik et al. (1994) to locate the human *WT1-AS* promoter. However, there were disadvantages to this strategy, namely, that control regulatory regions for *Wt1* could also have been knocked out. The promoter region of the human *WT1-AS* gene within intron 1 of *WT1* is within a domain adjacent to the WT1 antisense regulatory region (ARR) that is a preferentially methylated CpG island strongly implicated in the epigenetic regulation of the *WT1-AS* and *AWT1* transcripts, acting specifically as a *cis*-acting transcriptional silencer element of the *WT1-AS* and *AWT1* promoters (see section 1.4.13). Presuming the mouse *Wt1-AS* promoter is also within a region of epigenetic regulation, deletion of this region could directly alter expression of the *AWt1* gene.

5.5.2 Efficiency of molecular cloning procedures

As discussed in Chapter 4, with all multistep molecular cloning strategies much prior thought needs to go into construct design to allow for the most efficient number of cloning steps; this is because molecular cloning strategies are often complex and take many months if not years to complete. Although much thought went into the design of the targeting construct in order to minimise the number and difficultly of cloning steps, several alterations to the plan occurred during the generation of the construct. Only the cloning steps actually carried out have been discussed.

Importantly, it was found that using plasmids with different antibiotic resistance genes provided an increase in efficiency to clone vector and insert fragments, by preventing any remaining undigested, circular insert plasmids from growing following transformation. Although the technique of inactivating one antibiotic resistance gene and replacing with another was done using the Genome Priming System, it would have also been possible to sub-clone the DNA into a suitable plasmid containing the appropriate antibiotic resistance.

Another useful finding was that the efficiency of ligating an oligonucleotide linker into a plasmid can be increased by using blunt ended oligonucleotides. The sx3/sx4 linker was originally designed with *SalI* overhang ends but cloning proved unsuccessful. Upon redesigning the linker to have blunt ends, cloning was achieved.

5.5.3 Efficiency of gene targeting procedures

The efficiency of correct targeting following transfection of targeting constructs into ES cells varies depending on a number of factors, namely, whether co-isogenic DNA has been used, the size of the homologous arms, the positive and negative selection mechanisms and the method of transfection.

Following transfection of the MCRII parental targeting construct into ES cells it was found that 5 out of 266 geneticin resistant clones (<2%) were correctly targeted. This targeting efficiency was very low, especially compared to the targeting efficiency of the *Bsp1* targeting construct in Chapter 4 (83%) (see section 4.4.3). As co-isogenic DNA and electroporation were used in both situations these differences in efficiency can perhaps be explained by the length of homologous arms or the levels of selection used.

The homologous arms of the MCRII targeting construct were 762bp and ~3.2kb (overall homology ~4kb), whereas the homologous arms of the *Bsp1* targeting construct were ~4.5kb and ~3.4kb (overall homology ~7.9kb). As previously discussed (see section 4.5.2), the length of homology within a targeting construct is known to dramatically alter the targeting efficiency, but only within a certain range. It has been documented, for example, that when targeting the *hypoxanthine phosphoribosyltransferase* gene with homologous arms between 4.2kb and 6.0kb there was a 3-fold increase (1.6-fold per kb) in targeting efficiency, with the total length of homology being a more important factor than the length of the shorter arm of homology (Hasty et al., 1991). Interestingly, a separate study targeting the same gene showed that an increase in homology from 4kb to 9.1kb resulted in a ~20-fold increase in targeting efficiency (Thomas and Capecchi, 1987). The length of the homologous arms could therefore have

been a considerable factor affecting the difference in targeting efficiency between the MCRII and *Basp1* targeting constructs. Indeed, if the length of homology had been double for the MCRII targeting construct, perhaps an efficiency of as much as 40% would have been observed. The length of homology was not designed to be larger as this would have been difficult to achieve during the molecular cloning steps, due to the positioning of suitable restriction endonuclease sites.

As for the levels of selection used, whereas there were three levels of selection for the *Basp1* targeting construct: geneticin positive selection, diphtheria toxin negative selection and a splice acceptor allowing expression of the neomycin resistance gene only following integration downstream of an endogenous promoter, there were only two levels of selection used within the MCRII targeting construct: geneticin positive selection and diphtheria toxin negative selection. The third level of selection within the MCRII targeting construct was not an option due to requirement of the transcriptional terminator sequence used to truncate the *Wt1-AS* transcript and that there were no known splices in *Wt1-AS* prior to the MCRII region. The levels of selection were therefore the most likely factor affecting the difference in targeting efficiency between the two targeting constructs.

5.5.4 Efficiency of *cre*-loxP recombinations and the generation of mixed clones

After ES cell clones correctly targeted with the MCRII targeting construct had been identified, a second round of selection for *Cre*-loxP recombination events following transfection with *Cre* was carried out. Unlike the first round of selection whereby the neomycin resistance cassette acted as a positive selection mechanism and the DTA cassette acted as a negative selection mechanism, there was no selection mechanism for this second round of transfection. It was therefore expected that the efficiency of targeting would be very low.

Following transfection of C100 with *Cre* a high proportion of isolated clones had undergone loxP recombination events (85%). The first documented case using

the pCAG-CRE plasmid reported a maximum recombination frequency of 26% with no selection mechanism, although clones were assayed by observing the reporter activity of a linked gene (Araki et al., 1997). A similar recombination frequency (23.5%) was observed during *Cre*-loxP recombination events assayed by PCR of clones following targeting of a BAC transgene to the *Hprt* locus with no selection mechanism (Heaney et al., 2004). The 85% recombination frequency was therefore higher than expected.

A large number of the recombination events resulted in mixed clones (73%). This could have been due to the lack of a selection mechanism to kill most of the transfected cells resulting in independent clones growing too close to one another. This could have been resolved by plating out cells following transfection at a lower density than was used. Another possibility was that following picking of clones, *Cre*-loxP recombination events continued to occur within some of the cells of the cell line. For example, cells that contained a Δ1 deletion retain two loxP sites that would have been able to produce a Δ2 deletion, and cells that contained a Δ3 deletion retain two loxP sites that would have been able to produce a Δ2 deletion.

5.5.5 Loss of the 5' most loxP site

During screening of the ES cell clones transfected with the MCRII parental targeting construct, it was found that 4 out of the 5 correctly targeted clones were lacking the 5' most loxP site. The loss of the 5' most loxP site in targeted ES cell clones can be explained by homologous recombination taking place between the 3' homologous arm and the MCRII region, instead of the 3' and 5' homologous arms (Figure 5.2A). Although the MCRII region (and the short region of DNA sequence surrounding MCRII) within the targeting construct was only ~630bp, this length has been shown to be sufficient for homologous recombination to take place. Hasty et al. (1991) showed that as long as the overall length of homology was over 1.7kb, a short homologous arm of 472bp was as efficient as 1.2kb.

5.5.6 Efficiency of generating founder lines via blastocyst injection

The overall percentage of live births following blastocyst injection resulting in chimeric animals was 13%, and the overall percentage of those chimeras that generated founder lines due to the presence of genetically modified ES cells within the germline of the chimeras was 100% (2 out of 2 mice) (Table 5.4). For the *Baspl* knockout project (Chapter 4), these figures were 22% and 5% respectively. The discrepancy of 9% between the percentage of chimeras generated for each project is probably not high enough to be regarded as meaningful. It can perhaps be put down to the longer culture time of ES cells used to generate the MCRII targeted chimeric mice due to the two rounds of selection compared to the *Baspl* knockout chimeric mice resulting in more differentiated cells unable to colonize the blastocyst. Applying the same reasoning, the longer culture time might be expected to lower the germline transmission rates. As is evident, this was not the case. Perhaps the reasoning behind the higher rate of germline transmission for the MCRII targeted mice was a deleterious phenotype associated with the *Baspl* knockout mice.

5.5.7 Transcriptional termination of antisense transcripts

A key point that needs to be addressed is whether the transcriptional terminator is working to prematurely terminate *Wt1-AS* transcripts within the Travis line. As mentioned above, preliminary experiments were unable to resolve this question and further experiments are currently being carried out. Although this fact is still unknown, termination of antisense gene expression by the positioning of a transcriptional terminator sequence within the genomic locus has been shown to be successful in two mouse models employing similar strategies.

The first published example was the termination of the non-coding *Air* RNA at the *Igf2r* locus (Sleutels et al., 2002) (see section 1.3.1.3). The transcriptional terminator used in this case was a 1.2kb polyadenylation cassette from the β-globin gene. The second example of an antisense gene being transcriptionally terminated by the positioning of a transcription terminator sequence was the *Kcnq1ot1* transcript (*Lit1* transcript) at the *Kcnq1* locus (Mancini-DiNardo et al., 2006). The transcriptional terminator used in this case was a tandem of four

SV40 polyadenylation sequences. The transcriptional terminator used to truncate the *Wt1-AS* transcripts in this study was a tandem of three SV40 polyadenylation sequences known to efficiently block plasmid-initiated transcription in several systems when inserted upstream of a reporter gene (Maxwell et al., 1989) as well as transcriptional read-through from the eukaryotic promoter pGK within ES cells (Soriano, 1999).

5.5.8 Proteinuria analysis as a measure of kidney function

Urine was collected from Travis wildtype, heterozygous and homozygous mice at 4-5 months and 7-8 months of age and from wildtype mice and Travis homozygous mice at 1 year of age to test for the presence of protein. Compared to the ~66kDa positive control (10μg BSA), the bands in the positive lanes were faint. A recent study looking at proteinuria in transgenic mice with a single isoform of *Wt1* (WT1 KTS-) and a disrupted *Wt1* allele (Lahiri et al., 2007) ran 5μl of urine samples on an SDS-PAGE gel if the samples had a protein content over 20g/L as determined by proteinuria indicator strips; these 5μl samples contained at least 100μg of protein and appeared very bright on the gel following coomassie brilliant blue staining. None of the samples tested in this study gave bands of such strong staining intensity. Relatively low levels of protein (<10μg) were detected in a few animals but these represented all three genotypes. Most of the protein positive samples came from the oldest animals tested and whilst this might indicate mild proteinuria associated with age, there was no strong evidence that this was influenced by the MCRII targeted mutation. The only evidence suggestive of a genetic difference between the mice was a skew in the number of transgenic animals with protein positive samples. For example, in the 7-8 month old mice, 3 out of 20 wildtype mice (15%), 7 out of 21 heterozygous mice (33%) and 2 out of 15 homozygous mice (13%) had protein positive samples; and in the 1 year old mice, 1 out of 6 wildtype mice (16.7%) and 3 out of 6 homozygous mice (50%) had protein positive samples.

5.5.9 Body and organ weight analysis in the 7-8 month old cohort

Body and organ weights of wt, het and homo 7-8 month old Travis animals of both sexes have been analysed. In summary, the significant weight differences observed in the 7-8 month old weight study were:

- Female body weights - het being heavier than both wt and homo.
- Female relative brain weights - het brains being smaller than both wt and homo
- Male actual eye weights - wt being smaller than both het and homo
- Female relative eye weights - wt being larger than het

Any significant difference in weight of the het animals compared to the wt and homo animals, as was the case for female body weights and female relative brain weights, was difficult to explain. The prediction was that if any significant differences in weight were seen that they would be between wt and homo animals with het animals being perhaps like wt or an intermediate weight between wt and homo.

The differences between the male and female eye weights was also difficult to explain as there was no precedent for an organ being significantly larger in a mutant genotype in one sex but smaller in the opposite sex. It was possible that due to the small size of these organs and the accuracy of the balance used (3 decimal places), the eye weights were not large enough for them to be meaningfully compared.

Although a good starting point to pick up any obvious abnormalities, the weight study was secondary to the histological analysis of the organs. After all, a defect in organ morphology may not lead to a difference in size or weight if only the cellular structure had been altered.

5.5.10 Body and organ weight analysis in the 1 year old cohort

Body and organ weights of 1 year old male wt 129 animals and male homo Travis animals have been analysed. In summary, the significant weight differences observed in the 1 year old weight study were:

- Left and right testis weight (actual and relative) - homo being larger than wt

- Actual eye weights - homo being smaller than wt

As with the 7-8 month old cohort, it was possible that due to the small size of the eyes the weights were not accurate enough for them to be meaningfully compared. This is particularly likely to be the case as significance is inconsistent between the age groups (male wt eyes smaller at 7-8 months of age, male homo eyes smaller at 1 year of age).

The significant difference between the testis weights may be biologically relevant, as it could be an age related effect.

5.5.11 Differences in histology in the 7-8 month old cohort

Histological organ analysis has been carried out on the 7-8 month old Travis animals. Only two samples from each genotype for each sex were analyzed. This was due to the huge amount of work required to both section, stain and visualize the samples. Now that a potential difference has been found between the retina samples between the genotypes, a larger number of samples need to be screened. A study is currently taking place whereby the width of each cell layer of the retina is being measured and the number of cells counted for a larger cohort of animals.

5.5.12 Mice with morbid features or abnormal organs

As mentioned above, there were several Travis mice that became ill and had to be culled or were found dead in their cage. Although some of same morbid features did appear in more than one animal, the frequency of these phenotypes was not such that it was obviously due to their genotype. Indeed, the two cases of liver tumour appeared in one het mouse and one wt mouse. The two cases of abnormal testes were both within het mice, but they displayed very different morphologies.

5.5.13 What has been the effect of transcriptionally terminating the *Wt1-AS* transcript?

So far, no striking phenotype has been seen due to the transcriptional termination of *Wt1-AS* within the Travis line. The next step would be to look for genetic evidence of a change in Wt1 expression. This could be achieved by immunohistochemistry of organ sections, particularly the kidney, using a Wt1 antibody. Any difference in Wt1 expression levels could then be put down to *Wt1* and *Wt1-AS* being able to reciprocally regulate each others expression. Due to prior evidence of a positive feedback between the two transcripts, the hypothesis would be that a decrease in *Wt1-AS* expression due to transcriptional termination of the transcript would lead to a decrease in *Wt1* transcription and hence Wt1 protein. One point to bear in mind is that the truncation does not permit a direct examination of the role of the full length *Wt1-AS* transcript, only the role of the *Wt1-AS* transcript from and downstream of the MCRII region. Any differences within the Travis line would mean that MCRII functions at the level of RNA as part of *Wt1-AS* transcripts and could answer some of the questions posed within the introduction to this chapter, such as, does MCRII mediate regulation and distribution of *Wt1* sense RNA and Wt1 protein levels during mouse development?

5.5.14 Generation of the MCRII deletion line

Although targeted ES cell line C100 did not contain the 5' most loxP site and was therefore not capable of generating Δ2 deletions, there were four other targeted cell lines to use. Like C100, three other of the targeted lines (M101, C68 and C216) were also found to not contain the 5' most loxP site; C277 however did contain this site. Cell line C277 was used to generate Δ2 deletions and subsequently a mouse line named Davis. Although this mouse line was generated during the course of this study there was not time to carry out a phenotypical analysis and this is now being done by another member of the laboratory. Once completed, the two studies can be evaluated to answer all the questions posed within the introduction of this chapter and specifically to address whether MCRII functions at the level of DNA.

5.6 Appendix

mcr2 gtv#2 glycerol stock made on 24/2/04, maxiprep made on 25/2/04.

Final Discussion

6.1 Basp1 knockout mice

6.1.1 Role of Basp1

A mouse knockout model of *Basp1* has been generated from a single targeted ES cell line, however, a breeding line was not successfully established due to postnatal lethality of *Basp1* heterozygous knockout mice occurring around weaning age. The cause of this postnatal lethality is likely to be due to impaired digestion of food and nutrient uptake due to a physiological defect of the gut or due to behavioural reasons caused by a malformation within the brain, both regions of *Basp1* expression. This conclusion has been reached due to the observations that before death, mice appeared thin and no obvious abnormalities were detected upon necropsy other than the stomach and intestines being filled with air. The phenotype of the $Basp1^{+/-}$ mice described has similarities to that of $Basp1^{-/-}$ mice generated by Frey et al. (2000), and $Gap43^{-/-}$ mice generated by Strittmatter et al. (1995). $Basp1^{+/-}$ and $Basp1^{-/-}$ mice generated by Frey et al. (2000), were born in the expected Mendelian ratios. Approximately 66% of $Basp1^{-/-}$ mice died between birth and 4 months but only a small number of $Basp1^{+/-}$ mice died postnatally. The surviving $Basp1^{-/-}$ adults were 50-70% the weight of their wildtype littermates and the surviving $Basp1^{+/-}$ mice were 80% of normal weight. $Gap43^{+/-}$ and $Gap43^{-/-}$ mice generated by Strittmatter et al. (1995), were also born in the expected Mendelian ratios. 90% of the $Gap43^{-/-}$ mice died between birth and weaning, weight gain being poor and necropsy showing the stomach and gut were devoid of food.

The differences observed between the *Basp1* knockout mice discussed within this thesis and published by Frey et al.(2000) could be due to several possibilities. Firstly, the mice discussed within this thesis were designed to eliminate all Basp1 protein whereas insufficient detail was given about the design of the published knockout and it is therefore possible that the full length protein was not eliminated within this model, thereby reducing the severity of the phenotype. Secondly, it was possible that strain specific lethality could be caused by

modifier genes, however, it was not clear what strain background was used for the published knockout. Any differences in phenotypic severity between individuals could also be due to the action of these modifier genes. Thirdly, it was possible that an unidentified problem had occurred during the generation of the *Basp1* knockout mouse model discussed within this thesis, the postnatal lethality therefore not being due to the engineered mutation. To resolve this issue it would be desirable to generate further *Basp1* knockout mice from at least one additional, independently targeted ES cell line.

6.1.2 Interaction between Basp1 and Wt1

Within the adult mouse overlapping expression of *Basp1* and *Wt1* has only been observed in the kidney and testis (Armstrong et al., 1993; Carpenter et al., 2004). The primary aim of generating a *Basp1* knockout was to observe the effect of this mutation on the development of the kidney and the interaction of Basp1 and Wt1 within this organ in order to further our understanding of the genetics of Wilms' tumour. There were no obvious abnormalities within the kidneys of the $Basp1^{+/-}$ mice analysed, however these mice were only heterozygous for the mutation.

Although no phenotype was observed within the kidney, due to the haploinsuffiency of *Basp1*, other organs more sensitive to expression levels of Basp1 could have been affected. It remains possible that the interaction between these two proteins in other organs during an earlier developmental stage was the cause of the later postnatal lethality due to a defect and subsequent failure of certain organ systems over time. At E10-E10.5 for example there is a wide overlap of *Basp1* and *Wt1* expression within the mesothelial linings of certain organs contained within the coelomic cavity (see section 5.3.2).

Without a breeding line of *Basp1* knockout mice it was not possible to investigate the mechanism of action between Basp1 and Wt1. One hypothesis for the interaction between these two proteins is that in the absence of BASP1, the transcriptional activation domain of WT1 becomes constitutively active; BASP1 is thought to be a component of a WT1 cosuppressor that can aid regulation of the transcriptional activation domain of WT1 by binding to the suppression

domain of WT1 thereby blocking the function of the transcriptional activation domain. The result of a constitutively active transcriptional activation domain being that genes usually transcriptionally suppressed by WT1 are no longer suppressed, and may even be activated. Applying the same principle, reduced levels of BASP1 (as would be found in *Basp1* heterozygous knockout mice) might cause an increase in the activity of the transcriptional activation domain of WT1 and reduced suppression of genes usually transcriptionally suppressed by WT1. The overall result being loss of controlled gene expression within tissues expressing both BASP1 and WT1. In organs not expressing WT1 the mechanism of WT1-independent BASP1 action is unknown.

6.1.3 Future direction

As discussed above it would be desirable to generate further *Basp1* knockout mice from at least one additional, independently targeted ES cell line in order to compare the phenotype between the lines. If postnatal lethality was observed with this independent line, breeding the chimeras to mice of a different strain background may decrease the severity of the phenotype and prolong postnatal life. If a breeding knockout line cannot be established due to postnatal lethality, a *Basp1* conditional knockout mouse could be generated to specifically analyze *Basp1* deficiency within the kidney.

As *Basp1* is more widely expressed than *Wt1* during development, most notably within neurons, Basp1 must have a WT1-independent mode of action. In order to investigate this mode of action, it will be important to identify other proteins that can interact with Basp1. This could initially be addressed by a GST pull-down assay followed by 2-dimensional gel electrophoresis and mass spectrometry, a yeast two-hybrid screen or RNAi. As BASP1 is known to contain sites for myristylation and sumoylation, regulation of this protein is likely to be complex.

6.2 Wt1-AS truncation mouse mutant

6.2.1 Role of Wt1-AS

A breeding mouse line (Travis line) predicted to exhibit a truncation of the *Wt1-AS* transcript at the position of the MCRII sequence has been successfully generated. Heterozygous and homozygous mutant mice were compared to wildtypes in order to detect any phenotypic differences. An assay to detect the presence of protein within the urine was carried out along with an organ weight study and histological observations of organ sections; no obvious phenotypic abnormalities were seen although there was a subtle difference in the histology of the eye, such that the overall width of the retina layer in the homozygous samples appeared thinner than that of the wildtype and heterozygous samples.

6.2.2 Interaction between *Wt1-AS* and *Wt1*

The *WT1-AS* promoter has been shown to be transcriptionally activated by the full length WT1 KTS- isoform (Malik et al., 1995) although WT1 binding sites were not necessary for this transactivation, suggesting that DNA binding is not the mechanism by which WT1 exerts its affect on the *WT1-AS* promoter (Moorwood et al., 1999). *WT1-AS* is also known to positively modulate WT1 protein levels *in vitro*. The previous two points demonstrate positive feedback between *WT1-AS* and *WT1*. Although the mechanism whereby these two transcripts interact is unknown, presumably an absence of *WT1-AS* would cause the positive feedback loop between *WT1-AS* and *WT1* to break down, causing a decrease in *WT1* sense RNA and WT1 protein. It is possible that *WT1-AS* and WT1 exert their effects upon one another by forming RNA:RNA or RNA:protein duplexes. Indeed, it is known that RNA:RNA duplexes are able to form *in vivo* (Dallosso et al., 2007) and that RNA:protein duplexes are able to form *in vitro* (Ye et al., 1996); additionally *WT1-AS* is known to be exported from the nucleus, an essential condition if RNA:protein duplex formation are to occur *in vivo*. Mechanistic studies of the mouse line containing a truncation of the *Wt1-AS* transcript will be needed to find out what happens to Wt1 levels.

6.2.3 Future direction

As there are very few studies that have looked at the mechanistic role behind antisense transcription, and particularly as this is likely to be the first involving an antisense transcript acting independently of genomic imprinting, it is of scientific interest to pursue this study further.

Initially it will be important to ascertain if the transcriptional terminator within the Travis line is working to confirm that this mouse line is truly exhibiting a truncation of the *Wt1-AS* transcript.

A mechanistic study to look for genetic evidence of a change in Wt1 expression is still needed within the Travis line. This could be achieved by immunohistochemistry of organ sections using a Wt1 antibody to look for differences in expression patterns or by Western hybridization to detect differences in Wt1 expression levels. Further histological data within the eye also needs to be carried out to determine if there is truly a phenotypic difference between wildtypes and homozygous mutant mice. Additionally, analysis of the MCRII deletion line (Davis line) will complement the phenotypic data gathered from the Travis line in order to address the role of the *Wt1-AS* transcript and the MCRII DNA region.

References

Agranovsky, A. A. (1992). Exogenous Primer-Independent cDNA Synthesis with Commercial Reverse-Transcriptase Preparations on Plant-Virus RNA Templates. *Analytical Biochemistry* **203**, 163-165.

Aigner, L., Arber, S., Kapfhammer, J. P., Laux, T., Schneider, C., Botteri, F., Brenner, H. R. and Caroni, P. (1995). Overexpression of the Neural Growth-Associated Protein Gap-43 Induces Nerve Sprouting in the Adult Nervous-System of Transgenic Mice. *Cell* **83**, 269-278.

Araki, K., Araki, M., Miyazaki, J. and Vassalli, P. (1995). Site-Specific Recombination of a Transgene in Fertilized-Eggs by Transient Expression of Cre Recombinase. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 160-164.

Araki, K., Imaizumi, T., Okuyama, K., Oike, Y. and Yamamura, K. (1997). Efficiency of recombination by Cre transient expression in embryonic stem cells: Comparison of various promoters. *Journal of Biochemistry* **122**, 977-982.

Armstrong, J. F., Pritchard-Jones, K., Bickmore, W. A., Hastie, N. D. and Bard, J. B. L. (1993). The Expression of the Wilms-Tumor Gene, Wt1, in the Developing Mammalian Embryo. *Mechanisms of Development* **40**, 85-97.

Auerbach, W., Dunmore, J. H., Fairchild-Huntress, V., Fang, Q., Auerbach, A. B., Huszar, D. and Joyner, A. L. (2000). Establishment and chimera analysis of 129/SvEv- and C57BL/6- derived mouse embryonic stem cell lines. In *Biotechniques*, vol. 29, pp. 1024-+.

Blackshear, P. J. (1993). The Marcks Family of Cellular Protein-Kinase-C Substrates. *Journal of Biological Chemistry* **268**, 1501-1504.

Bonetta, L., Kuehn, S. E., Huang, A., Law, D. J., Kalikin, L. M., Koi, M., Reeve, A. E., Brownstein, B. H., Yeger, H., Williams, B. R. G. et al. (1990). Wilms-Tumor Locus on 11p13 Defined by Multiple Cpg Island- Associated Transcripts. *Science* **250**, 994-997.

Brightwell, G., Poirier, V., Cole, E., Ivins, S. and Brown, K. W. (1997). Serum-dependent and cell cycle-dependent expression from a cytomegalovirus-based mammalian expression vector. *Gene* **194**, 115-123.

Brown, K. W. and Malik, K. T. (2001). The molecular biology of Wilms tumour. *Expert Rev Mol Med* **2001**, 1-16.

Bruening, W. and Pelletier, J. (1996). A non-AUG translational initiation event generates novel WT1 isoforms. *Journal of Biological Chemistry* **271**, 8646-8654.

Buckler, A. J., Pelletier, J., Haber, D. A., Glaser, T. and Housman, D. E. (1991). Isolation, Characterization, and Expression of the Murine Wilms-Tumor Gene (Wt1) During Kidney Development. *Molecular and Cellular Biology* **11**, 1707-1712.

Call, K. M., Glaser, T., Ito, C. Y., Buckler, A. J., Pelletier, J., Haber, D. A., Rose, E. A., Kral, A., Yeger, H., Lewis, W. H. et al. (1990). Isolation and Characterization of a Zinc Finger Polypeptide Gene at the Human Chromosome-11 Wilms Tumor Locus. *Cell* **60**, 509-520.

Campbell, C. E., Huang, A., Gurney, A. L., Kessler, P. M., Hewitt, J. A. and Williams, B. R. G. (1994). Antisense Transcripts and Protein-Binding Motifs within the Wilms-Tumor (Wt1) Locus. *Oncogene* **9**, 583-595.

Caricasole, A., Duarte, A., Larsson, S. H., Hastie, N. D., Little, M., Holmes, G., Todorov, I. and Ward, A. (1996). RNA binding by the Wilms tumor suppressor zinc finger proteins. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 7562-7566.

Caroni, P. (1997). Overexpression of growth-associated proteins in the neurons of adult transgenic mice. *Journal of Neuroscience Methods* **71**, 3-9.

Caroni, P., Aigner, L. and Schneider, C. (1997). Intrinsic neuronal determinants locally regulate extrasynaptic and synaptic growth at the adult neuromuscular junction. *Journal of Cell Biology* **136**, 679-692.

Carpenter, B., Hill, K. J., Charalambous, M., Wagner, K. J., Lahiri, D., James, D. I., Andersen, J. S., Schumacher, V., Royer-Pokora, B., Mann, M. et al. (2004). BASP1 is a transcriptional cosuppressor for the Wilms' tumor suppressor protein WT1. *Molecular and Cellular Biology* **24**, 537-549.

Coggins, P. J. and Zwiers, H. (1991). B-50 (Gap-43) - Biochemistry and Functional Neurochemistry of a Neuron-Specific Phosphoprotein. *Journal of Neurochemistry* **56**, 1095-1106.

Costa, F. F. (2007). Non-coding RNAs: Lost in translation? *Gene* **386**, 1-10.

Dallosso, A. R., Hancock, A. L., Brown, K. W., Williams, A. C., Jackson, S. and Malik, K. (2004). Genomic imprinting at the WT1 gene involves a novel coding transcript (AWT1) that shows deregulation in Wilms' tumours. *Human Molecular Genetics* **13**, 405-415.

Dallosso, A. R., Hancock, A. L., Malik, S., Salpekar, A., King-Underwood, L., Pritchard-Jones, K., Peters, J., Moorwood, K., Ward, A., Malik, K. T. A. et al. (2007). Alternately spliced *WT1* antisense transcripts interact with *WT1* sense RNA and show epigenetic and splicing defects in cancer. *RNA* **13**, 1-13.

Davies, R. C., Calvio, C., Bratt, E., Larsson, S. H., Lamond, A. I. and Hastie, N. D. (1998). WT1 interacts with the splicing factor U2AF65 in an isoform-dependent manner and can be incorporated into spliceosomes. *Genes & Development* **12**, 3217-3225.

Discenza, M. T., He, S. J., Lee, T. H., Chu, L. L., Bolon, B., Goodyer, P., Eccles, M. and Pelletier, J. (2003). WT1 is a modifier of the Pax2 mutant phenotype: cooperation and interaction between WT1 and Pax2. *Oncogene* **22**, 8145-8155.

Discenza, M. T. and Pelletier, J. (2004). Insights into the physiological role of WT1 from studies of genetically modified mice. *Physiological Genomics* **16**, 287-300.

Doetschman, T. (1999). Interpretation of phenotype in genetically engineered mice. *Laboratory Animal Science* **49**, 137-143.

Eccles, M. R., Grubb, G., Ogawa, O., Szeto, J. and Reeve, A. E. (1994). Cloning of Novel Wilms-Tumor Gene (Wt1) Cdnas - Evidence for Antisense Transcription of Wt1. *Oncogene* **9**, 2059-2063.

Eggan, K., Rode, A., Jentsch, I., Samuel, C., Hennek, T., Tintrup, H., Zevnik, B., Erwin, J., Loring, J., Jackson-Grusby, L. et al. (2002). Male and female mice derived from the same embryonic stem cell clone by tetraploid embryo complementation. *Nature Biotechnology* **20**, 455-459.

Eggenschwiler, J., Ludwig, T., Fisher, P., Leighton, P. A., Tilghman, S. M. and Efstratiadis, A. (1997). Mouse mutant embryos overexpressing IGF-II exhibit phenotypic features of the Beckwith-Wiedemann and Simpson-Golabi-Behmel syndromes. *Genes & Development* **11**, 3128-3142.

Englert, C., Hou, X., Maheswaran, S., Bennett, P., Ngwu, C., Re, G. G., Garvin, A. J., Rosner, M. R. and Haber, D. A. (1995). Wt1 Suppresses Synthesis of the Epidermal Growth-Factor Receptor and Induces Apoptosis. *Embo Journal* **14**, 4662-4675.

Favor, J., Sandulache, R., NeuhauserKlaus, A., Pretsch, W., Chatterjee, B., Senft, E., Wurst, W., Blanquet, V., Grimes, P., Sporle, R. et al. (1996). The mouse Pax2(1Neu) mutation is identical to a human PAX2 mutation in a family with renal-coloboma syndrome and results in developmental defects of the brain, ear, eye, and kidney. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 13870-13875.

Fedorov, L. M., HaegelKronenberger, H. and Hirchenhain, J. (1997). A comparison of the germline potential of differently aged ES cell lines and their transfected descendants. *Transgenic Research* **6**, 223-231.

Fraizer, G. C., Patmasiriwat, P., Zhang, X. H. and Saunders, G. F. (1995). Expression of the Tumor-Suppressor Gene Wt1 in Both Human and Mouse Bone-Marrow. *Blood* **86**, 4704-4706.

Frey, D., Laux, T., Xu, L., Schneider, C. and Caroni, P. (2000). Shared and unique roles of CAP23 and GAP43 in actin regulation, neurite outgrowth, and anatomical plasticity. *Journal of Cell Biology* **149**, 1443-1453.

Friedrich, G. and Soriano, P. (1991). Promoter Traps in Embryonic Stem-Cells - a Genetic Screen to Identify and Mutate Developmental Genes in Mice. *Genes & Development* **5**, 1513-1523.

Gamsjaeger, R., Liew, C. K., Loughlin, F. E., Crossley, M. and Mackay, J. P. (2007). Sticky fingers: zinc-fingers as protein-recognition motifs. *Trends in Biochemical Sciences* **32**, 63-70.

Gao, F., Maiti, S., Alam, N., Zhang, Z., Deng, J. M., Behringer, R. R., Lecureuil, C., Guillou, F. and Huff, V. (2006). The Wilms tumor gene, Wt1, is required for Sox9 expression and maintenance of tubular architecture in the developing testis. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 11987-11992.

Gao, F., Maiti, S., Sun, G. Z., Ordonez, N. G., Udtha, M., Deng, J. M., Behringer, R. R. and Huff, V. (2004). The Wt1(+/R394W) mouse displays glomerulosclerosis and early-onset renal failure characteristic of human Denys-Drash syndrome. *Molecular and Cellular Biology* **24**, 9899-9910.

Gessler, M. and Bruns, G. A. P. (1993). Sequence of the Wt1 Upstream Region Including the Wit-1 Gene. *Genomics* **17**, 499-501.

Gessler, M., Konig, A. and Bruns, G. A. P. (1992). The Genomic Organization and Expression of the Wt1 Gene. *Genomics* **12**, 807-813.

Gessler, M., Poustka, A., Cavenee, W., Neve, R. L., Orkin, S. H. and Bruns, G. A. P. (1990). Homozygous Deletion in Wilms-Tumours of a Zinc-Finger Gene Identified by Chromosome Jumping. *Nature* **343**, 774-778.

Gilbert, S. F. (2000). Developmental Biology - Sixth Edition.

Gill, G. (2004). SUMO and ubiquitin in the nucleus: different functions, similar mechanisms? *Genes & Development* **18**, 2046-2059.

Glaser, T., Lane, J. and Housman, D. (1990). A Mouse Model of the Aniridia-Wilms Tumor Deletion Syndrome. *Science* **250**, 823-827.

Gong, Y. L., Eggert, H. and Englert, C. (2001). The murine Wilms tumor suppressor gene (wt1) locus. *Gene* **279**, 119-126.

Guo, J. K., Menke, A. L., Gubler, M. C., Clarke, A. R., Harrison, D., Hammes, A., Hastie, N. D. and Schedl, A. (2002). WT1 is a key regulator of podocyte function: reduced expression levels cause crescentic glomerulonephritis and mesangial sclerosis. *Human Molecular Genetics* **11**, 651-659.

Haber, D. A., Park, S., Maheswaran, S., Englert, C., Re, G. G., Hazenmartin, D. J., Sens, D. A. and Garvin, A. J. (1993). Wt1-Mediated Growth Suppression of Wilms-Tumor Cells Expressing a Wt1 Splicing Variant. *Science* **262**, 2057-2059.

Haber, D. A., Sohn, R. L., Buckler, A. J., Pelletier, J., Call, K. M. and Housman, D. E. (1991). Alternative Splicing and Genomic Structure of the Wilms-Tumor Gene-Wt1. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 9618-9622.

Halford, M. M., Armes, J., Buchert, M., Meskenaite, V., Grail, D., Hibbs, M. L., Wilks, A. F., Farlie, P. G., Newgreen, D. F., Hovens, C. M. et al. (2000). Ryk-deficient mice exhibit craniofacial defects associated with perturbed Eph receptor crosstalk. *Nature Genetics* **25**, 414-418.

Hammes, A., Guo, J. K., Lutsch, G., Leheste, J. R., Landrock, D., Ziegler, U., Gubler, M. C. and Schedl, A. (2001). Two splice variants of the Wilms' tumor 1 gene have distinct functions during sex determination and nephron formation. *Cell* **106**, 319-329.

Hancock, A. L., Brown, K. W., Moorwood, K., Moon, H., Holmgren, C., Mardikar, S. H., Dallosso, A. R., Klenova, E., Loukinov, D., Ohlsson, R. et al. (2007). A CTCF-binding silencer regulates the imprinted genes AWT1 and WT1-AS and exhibits sequential epigenetic defects during Wilms' tumourigenesis. *Human Molecular Genetics* **16**, 343-354.

Hastie, N. D. (1994). The Genetics of Wilms-Tumor - a Case of Disrupted Development. *Annual Review of Genetics* **28**, 523-558.

Hasty, P., Riveraperez, J. and Bradley, A. (1991). The Length of Homology Required for Gene Targeting in Embryonic Stem-Cells. *Molecular and Cellular Biology* **11**, 5586-5591.

Heaney, J. D., Rettew, A. N. and Bronson, S. K. (2004). Tissue-specific expression of a BAC transgene targeted to the Hprt locus in mouse embryonic stem cells. *Genomics* **83**, 1072-1082.

Herzer, U., Crocoll, A., Barton, D., Howells, N. and Englert, C. (1999). The Wilms tumor suppressor gene wt1 is required for development of the spleen. *Current Biology* **9**, 837-840.

Hohenstein, P. and Hastie, N. D. (2006). The many facets of the Wilms' tumour gene, WT1. *Human Molecular Genetics* **15**, R196-R201.

Huang, A., Campbell, C. E., Bonetta, L., McAndrewshill, M. S., Chiltonmacneill, S., Coppes, M. J., Law, D. J., Feinberg, A. P., Yeger, H. and Williams, B. R. G. (1990). Tissue, Developmental, and Tumor-Specific Expression of Divergent Transcripts in Wilms-Tumor. *Science* **250**, 991-994.

Iino, S., Kobayashi, S. and Maekawa, S. (1999). Immunohistochemical localization of a novel acidic calmodulin- binding protein, NAP-22, in the rat brain. *Neuroscience* **91**, 1435-1444.

Iino, S. and Maekawa, S. (1999). Immunohistochemical demonstration of a neuronal calmodulin- binding protein, NAP-22, in the rat spinal cord. *Brain Research* **834**, 66-73.

Iino, S., Taguchi, K., Maekawa, S. and Nojyo, Y. (2004). Motor, sensory and autonomic nerve terminals containing NAP-22 immunoreactivity in the rat muscle. *Brain Research* **1002**, 142-150.

Jinno, Y., Yun, K. K., Nishiwaki, K., Kubota, T., Ogawa, O., Reeve, A. E. and Niikawa, N. (1994). Mosaic and Polymorphic Imprinting of the Wt1 Gene in Humans. *Nature Genetics* **6**, 305-309.

Kaneuchi, M., Sasaki, M., Tanaka, Y., Shiina, H., Yamada, H., Yamamoto, R., Sakuragi, N., Enokida, H., Verma, M. and Dahiya, R. (2005). WT1 and WT1-AS genes are inactivated by promoter methylation in ovarian clear cell adenocarcinoma. *Cancer* **104**, 1924-1930.

King-Underwood, L. and Pritchard-Jones, K. (1998). Wilms' tumor (WT1) gene mutations occur mainly in acute myeloid leukemia and may confer drug resistance. *Blood* **91**, 2961-2968.

Kiyosawa, H., Yamanaka, I., Osato, N., Kondo, S. and Hayashizaki, Y. (2003). Antisense transcripts with FANTOM2 clone set and their implications for gene regulation. *Genome Research* **13**, 1324-1334.

Knudson, A. G. (1971). Mutation and cancer: Statistical study of Retinoblastoma. *Proceedings of the National Academy of Sciences of the United States of America* **68**, 820-823.

Kreidberg, J. A., Natoli, T. A., McGinnis, L., Donovan, M., Biggers, J. D. and Amstutz, A. (1999). Coordinate action of Wt1 and a modifier gene supports embryonic survival in the oviduct. *Molecular Reproduction and Development* **52**, 366-375.

Kreidberg, J. A., Sariola, H., Loring, J. M., Maeda, M., Pelletier, J., Housman, D. and Jaenisch, R. (1993). Wt-1 Is Required for Early Kidney Development. *Cell* **74**, 679-691.

Lahiri, D., Dutton, J. R., Duarte, A., Moorwood, K., Graham, C. F. and Ward, A. (2007). Nephropathy and defective spermatogenesis in mice transgenic for a single isoform of the Wilms' tumour suppressor protein, WT1-KTS, together with one disrupted Wt1 allele. *Molecular Reproduction and Development* **74**, 300-311.

Larsson, S. H., Charlieu, J. P., Miyagawa, K., Engelkamp, D., Rassoulzadegan, M., Ross, A., Cuzin, F., Vanheyningen, V. and Hastie, N. D. (1995). Subnuclear Localization of Wt1 in Splicing or Transcription Factor Domains Is Regulated by Alternative Splicing. *Cell* **81**, 391-401.

Latchman, D. S. (1997). Transcription factors: An overview. *International Journal of Biochemistry & Cell Biology* **29**, 1305-1312.

Laux, T., Fukami, K., Thelen, M., Golub, T., Frey, D. and Caroni, P. (2000). GAP43, MARCKS, and CAP23 modulate PI(4,5)P-2 at plasmalemmal rafts, and regulate cell cortex actin dynamics through a common mechanism. *Journal of Cell Biology* **149**, 1455-1471.

Lecureuil, C., Fontaine, I., Crepieux, P. and Guillou, F. (2002). Sertoli and granulosa cell-specific Cre recombinase activity in transgenic mice. *Genesis* **33**, 114-118.

Ledermann, B. and Burki, K. (1991). Establishment of a Germ-Line Competent C57bl/6 Embryonic Stem- Cell Line. *Experimental Cell Research* **197**, 254-258.

Little, M. H., Dunn, R., Byrne, J. A., Seawright, A., Smith, P. J., Pritchardjones, K., Vanheyningen, V. and Hastie, N. D. (1992). Equivalent Expression of Paternally and Maternally Inherited Wt1 Alleles in Normal Fetal Tissue and Wilms-Tumors. *Oncogene* **7**, 635-641.

Longo, L., Bygrave, A., Grosveld, F. G. and Pandolfi, P. P. (1997). The chromosome make-up of mouse embryonic stem cells is predictive of somatic and germ cell chimaerism. *Transgenic Research* **6**, 321-328.

Maekawa, S., Maekawa, M., Hattori, S. and Nakamura, S. (1993). Purification and Molecular-Cloning of a Novel Acidic Calmodulin-Binding Protein from Rat-Brain. *Journal of Biological Chemistry* **268**, 13703-13709.

Maekawa, S., Murofushi, H. and Nakamura, S. (1994). Inhibitory Effect of Calmodulin on Phosphorylation of Nap-22 with Protein-Kinase-C. *Journal of Biological Chemistry* **269**, 19462-19465.

Maekawa, S., Sato, C., Kitajima, K., Funatsu, N., Kumanogoh, H. and Sokawa, Y. (1999). Cholesterol-dependent localization of NAP-22 on a neuronal membrane microdomain (Raft). *Journal of Biological Chemistry* **274**, 21369-21374.

Maheswaran, S., Park, S., Bernard, A., Morris, J. F., Rauscher, F. J., Hill, D. E. and Haber, D. A. (1993). Physical and Functional Interaction between Wt1 and P53 Proteins. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 5100-5104.

Malik, K., Salpekar, A., Hancock, A., Moorwood, K., Jackson, S., Charles, A. and Brown, K. W. (2000). Identification of differential methylation of the WT1 antisense regulatory region and relaxation of imprinting in Wilms' tumor. *Cancer Research* **60**, 2356-2360.

Malik, K., Yan, P., Huang, T. H. M. and Brown, K. W. (2001). Wilms' tumor: A paradigm for the new genetics. *Oncology Research* **12**, 441-449.

Malik, K. T. A., Poirier, V., Ivins, S. M. and Brown, K. W. (1994). Autoregulation of the Human Wt1 Gene Promoter. *Febs Letters* **349**, 75-78.

Malik, K. T. A., Wallace, J. I., Ivins, S. M. and Brown, K. W. (1995). Identification of an Antisense Wt1 Promoter in Intron-1 - Implications for Wt1 Gene-Regulation. *Oncogene* **11**, 1589-1595.

Mancini-DiNardo, D., Steele, S. J. S., Levorse, J. M., Ingram, R. S. and Tilghman, S. M. (2006). Elongation of the Kcnq1ot1 transcript is required for genomic imprinting of neighboring genes. *Genes & Development* **20**, 1268-1282.

Martini, F. H., Timmons, M. J. and Tallitsch, B. (2002). Human Anatomy - Fourth Edition: Benjamin Cummings.

Maston, G. A., Evans, S. K. and Green, M. R. (2006). Transcriptional regulatory elements in the human genome. *Annual Review of Genomics and Human Genetics* **7**, 29-59.

Matsunaga, E. (1981). Genetics of Wilms Tumor. *Human Genetics* **57**, 231-246.

Maxwell, I. H., Harrison, G. S., Wood, W. M. and Maxwell, F. (1989). A DNA Cassette Containing a Trimerized Sv40 Polyadenylation Signal Which Efficiently Blocks Spurious Plasmid-Initiated Transcription. *Biotechniques* **7**, 276-&.

McKay, L. M., Carpenter, B. and Roberts, S. G. E. (1999). Regulation of the Wilms' tumour suppressor protein transcriptional activation domain. *Oncogene* **18**, 6546-6554.

Menke, A., McInnes, L., Hastie, N. D. and Schedl, A. (1998a). The Wilms' tumor suppressor WT1: Approaches to gene function. *Kidney International* **53**, 1512-1518.

Menke, A. L., Clarke, A. R., Leitch, A., Ijpenberg, A., Williamson, K. A., Spraggon, L., Harrison, D. J. and Hastie, N. D. (2002). Genetic interactions between the Wilms' Tumor 1 gene and the p53 gene. *Cancer Research* **62**, 6615-6620.

Menke, A. L., Ijpenberg, A., Fleming, S., Ross, A., Medine, C. N., Patek, C. E., Spraggon, L., Hughes, J., Clarke, A. R. and Hastie, N. D. (2003). The wtl-heterozygous mouse; a model to study the development of glomerular sclerosis. *Journal of Pathology* **200**, 667-674.

Menke, A. L., van der Eb, A. J. and Jochemsen, A. G. (1998b). The Wilms' tumor 1 gene: Oncogene or tumor suppressor gene? In *International Review of Cytology - a Survey of Cell Biology, Vol 181*, vol. 181, pp. 151-212.

Miles, C. G., Slight, J., Spraggon, L., O'Sullivan, M., Patek, C. and Hastie, N. D. (2003). Mice lacking the 68-amino-acid, mammal-specific N-terminal extension of WT1 develop normally and are fertile. *Molecular and Cellular Biology* **23**, 2608-2613.

Moore, A. W., McInnes, L., Kreidberg, J., Hastie, N. D. and Schedl, A. (1999). YAC complementation shows a requirement for Wt1 in the development of epicardium, adrenal gland and throughout nephrogenesis. *Development* **126**, 1845-1857.

Moore, A. W., Schedl, A., McInnes, L., Doyle, M., Hecksher-Sorensen, J. and Hastie, N. D. (1998). YAC transgenic analysis reveals Wilms' Tumour 1 gene activity in the proliferating coelomic epithelium, developing diaphragm and limb. *Mechanisms of Development* **79**, 169-184.

Moorwood, K., Charles, A. K., Salpekar, A., Wallace, J. I., Brown, K. W. and Malik, K. (1998). Antisense WT1 transcription parallels sense mRNA and protein expression in fetal kidney and can elevate protein levels in vitro. *Journal of Pathology* **185**, 352-359.

Moorwood, K., Salpekar, A., Ivins, S. M., Hall, J., Powlesland, R. M., Brown, K. W. and Malik, K. (1999). Transactivation of the WT1 antisense promoter is unique to the WT1 +/- isoform. *Febs Letters* **456**, 131-136.

Moreira, R. F. and Noren, C. J. (1995). Minimum Duplex Requirements for Restriction Enzyme Cleavage near the Termini of Linear DNA Fragments. *Biotechniques* **19**, 56-59.

Morris, J. S., Stein, T., Pringle, M. A., Davies, C. R., Weber-Hall, S., Ferrier, R. K., Bell, A. K., Heath, V. J. and Gusterson, B. A. (2006). Involvement of axonal guidance proteins and their signaling partners in the developing mouse mammary gland. *Journal of Cellular Physiology* **206**, 16-24.

Mosevitsky, M. I., Caponi, J. P., Skladchikova, G. Y., Novitskaya, V. A. and Plekhanov, A. Y. (1996). Specific characteristics and primary structure of protein BASP1 initially found in axon terminals of neurons. *Biochemistry-Moscow* **61**, 864-871.

Mosevitsky, M. I., Capony, J. P., Skladchikova, G. Y., Novitskaya, V. A., Plekhanov, A. Y. and Zakharov, V. V. (1997). The BASP1 family of myristoylated proteins abundant in axonal termini. Primary structure analysis and physico-chemical properties. *Biochimie* **79**, 373-384.

Mosevitsky, M. I., Novitskaya, V. A., Plekhanov, A. Y. and Skladchikova, G. Y. (1994). Neuronal Protein Gap-43 Is a Member of Novel Group of Brain Acid-Soluble Proteins (Basps). *Neuroscience Research* **19**, 223-228.

Muller, U. (1999). Ten years of gene targeting: targeted mouse mutants, from vector design to phenotype analysis. *Mechanisms of Development* **82**, 3-21.

Mundel, P., Reiser, J., Borja, A. Z. M., Pavenstadt, H., Davidson, G. R., Kriz, W. and Zeller, R. (1997). Rearrangements of the cytoskeleton and cell contacts induce process formation during differentiation of conditionally immortalized mouse podocyte cell lines. *Experimental Cell Research* **236**, 248-258.

Mundlos, S., Pelletier, J., Darveau, A., Bachmann, M., Winterpacht, A. and Zabel, B. (1993). Nuclear-Localization of the Protein Encoded by the Wilms-Tumor Gene Wt1 in Embryonic and Adult Tissues. *Development* **119**, 1329-1341.

Nagy, A., Gertsenstein, M., Vintersten, K. and Behringer, R. (2003). Manipulating the mouse embryo - A Laboratory Manual: Spring Habour Laboratory Press.

Natoli, T. A., Liu, J., Eremina, V., Hodgens, K., Li, C., Hamano, Y., Mundel, P., Kalluri, R., Miner, J. H., Quaggin, S. E. et al. (2002a). A mutant form of the Wilms' tumor suppressor gene WT1 observed in Denys-Drash syndrome interferes with glomerular capillary development. *Journal of the American Society of Nephrology* **13**, 2058-2067.

Natoli, T. A., McDonald, A., Alberta, J. A., Taglienti, M. E., Housman, D. E. and Kreidberg, J. A. (2002b). A mammal-specific Exon of WT1 is not required for development or fertility. *Molecular and Cellular Biology* **22**, 4433-4438.

Niksic, M., Slight, J., Sanford, J. R., Caceres, J. F. and Hastie, N. D. (2004). The Wilms' tumour protein (WT1) shuttles between nucleus and cytoplasm and is present in functional polysomes. *Human Molecular Genetics* **13**, 463-471.

Nomura, S., Hogan, B. L. M., Wills, A. J., Heath, J. K. and Edwards, D. R. (1989). Developmental Expression of Tissue Inhibitor of Metalloproteinase (Timp) Rna. *Development* **105**, 575-583.

Novitskaya, V. A., Skladchikova, G. Y., Plekhanov, A. Y. and Mosevitskii, M. I. (1994). Detection of Brain Protein Basp1 in Rat Reproductive Tissue. *Doklady Akademii Nauk* **335**, 101-102.

O'Neill, M. J. (2005). The influence of non-coding RNAs on allele-specific gene expression in mammals. *Human Molecular Genetics* **14**, R113-R120.

Osoegawa, K., Tateno, M., Woon, P. Y., Frengen, E., Mammoser, A. G., Catanese, J. J., Hayashizaki, Y. and de Jong, P. J. (2000). Bacterial artificial chromosome libraries for mouse sequencing and functional analysis. *Genome Research* **10**, 116-128.

Park, S., Schalling, M., Bernard, A., Maheswaran, S., Shipley, G. C., Roberts, D., Fletcher, J., Shipman, R., Rheinwald, J., Demetri, G. et al. (1993). The Wilms-Tumor Gene Wt1 Is Expressed in Murine Mesoderm-Derived Tissues and Mutated in a Human Mesothelioma. *Nature Genetics* **4**, 415-420.

Park, S. Y., Kim, Y. I., Kim, B. G., Seong, C. S., Oh, Y. S., Baek, K. G. and Yoon, J. S. (1998). Characterization of bovine and human cDNAs encoding NAP-22 (22 kDa neuronal tissue-enriched acidic protein) homologs. *Molecules and Cells* **8**, 471-477.

Patek, C. E., Brownstein, D. G., Fleming, S., Wroe, C., Rose, L., Webb, A., Berry, R. L., Devenney, P. S., Walker, M., Maddocks, O. D. K. et al. (2007). Effects on kidney disease, fertility and development in mice inheriting a protein-truncating Denys-Drash syndrome allele (Wt1tmT396). *Transgenic Research DOI:10.1007/s11248-007-9157-0.*

Patek, C. E., Little, M. H., Fleming, S., Miles, C., Charlieu, J. P., Clarke, A. R., Miyagawa, K., Christie, S., Doig, J., Harrison, D. J. et al. (1999). A zinc finger truncation of murine WT1 results in the characteristic urogenital abnormalities of Denys-Drash syndrome. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 2931-2936.

Patterson, L. T., Pembaur, M. and Potter, S. S. (2001). Hoxa11 and Hoxd11 regulate branching morphogenesis of the ureteric bud in the developing kidney. *Development* **128**, 2153-2161.

Pearson, W. R. (1991). Searching Protein-Sequence Libraries - Comparison of the Sensitivity and Selectivity of the Smith-Waterman and Fasta Algorithms. *Genomics* **11**, 635-650.

Pelletier, J., Schalling, M., Buckler, A. J., Rogers, A., Haber, D. A. and Housman, D. (1991). Expression of the Wilms-Tumor Gene Wt1 in the Murine Urogenital System. *Genes & Development* **5**, 1345-1356.

Pritchard-Jones, K., Fleming, S., Davidson, D., Bickmore, W., Porteous, D., Gosden, C., Bard, J., Buckler, A., Pelletier, J., Housman, D. et al. (1990). The Candidate Wilms-Tumor Gene Is Involved in Genitourinary Development. *Nature* **346**, 194-197.

Rackley, R. R., Flenniken, A. M., Kuriyan, N. P., Kessler, P. M., Stoler, M. H. and Williams, B. R. G. (1993). Expression of the Wilms-Tumor Suppressor Gene Wt1 During Mouse Embryogenesis. *Cell Growth & Differentiation* **4**, 1023-1031.

Ravasi, T., Suzuki, H., Pang, K. C., Katayama, S., Furuno, M., Okunishi, R., Fukuda, S., Ru, K. L., Frith, M. C., Gongora, M. M. et al. (2006). Experimental validation of the regulated expression of large numbers of non-coding RNAs from the mouse genome. *Genome Research* **16**, 11-19.

Reddy, J. C. and Licht, J. D. (1996). The WT1 Wilms' tumor suppressor gene: How much do we really know? *Biochimica Et Biophysica Acta-Reviews on Cancer* **1287**, 1-28.

Riele, H. T., Maandag, E. R. and Berns, A. (1992). Highly Efficient Gene Targeting in Embryonic Stem-Cells through Homologous Recombination with Isogenic DNA Constructs. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 5128-5132.

Rivera, M. N. and Haber, D. A. (2005). Wilms' tumour: Connecting tumorigenesis and organ development in the kidney. *Nature Reviews Cancer* **5**, 699-712.

Rivera, M. N., Kim, W. J., Wells, J., Driscoll, D. R., Brannigan, B. W., Han, M., Kim, J. C., Feinberg, A. P., Gerald, W. L., Vargas, S. O. et al. (2007). An X chromosome gene, WTX, is commonly inactivated in Wilms tumor. *Science* **315**, 642-645.

Rucker, E. B., Dierisseau, P., Wagner, K. U., Garrett, L., Wynshaw-Boris, A., Flaws, J. A. and Hennighausen, L. (2000). Bcl-x and bax regulate mouse primordial germ cell survival and apoptosis during embryogenesis. *Molecular Endocrinology* **14**, 1038-1052.

Rupprecht, H. D., Drummond, I. A., Madden, S. L., Rauscher, F. J. and Sukhatme, V. P. (1994). The Wilms-Tumor Suppressor Gene Wt1 Is Negatively Autoregulated. *Journal of Biological Chemistry* **269**, 6198-6206.

Sadhu, A. (2006). Thesis Title: Role of the neuronal protein Cap23 in the maturation and maintenance of dendritic arbors in-vivo. In *Friedrich Miescher Institute*, (ed.: University of Basel.

Sainio, K., Hellstedt, P., Kreidberg, J. A., Saxen, L. and Sariola, H. (1997). Differential regulation of two sets of mesonephric tubules by WT-1. *Development* **124**, 1293-1299.

Salaun, C., James, D. J. and Chamberlain, L. H. (2004). Lipid rafts and the regulation of exocytosis. *Traffic* **5**, 255-264.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). Molecular Cloning - A Laboratory Manual: Cold Spring Harbour Laboratory Press.

Scharnhorst, V., Dekker, P., van der Eb, A. J. and Jochemsen, A. G. (1999). Internal translation initiation generates novel WT1 protein isoforms with distinct biological properties. *Journal of Biological Chemistry* **274**, 23456-23462.

Scharnhorst, V., van der Eb, A. J. and Jochemsen, A. G. (2001). WT1 proteins: functions in growth and differentiation. *Gene* **273**, 141-161.

Scholz, H., Bossone, S. A., Cohen, H. T., Akella, U., Strauss, W. M. and Sukhatme, V. P. (1997). A far upstream cis-element is required for Wilms' tumor-1 (WT1) gene expression in renal cell culture. *Journal of Biological Chemistry* **272**, 32836-32846.

Schroeder, W. T., Chao, L. Y., Dao, D. D., Strong, L. C., Pathak, S., Riccardi, V., Lewis, W. H. and Saunders, G. F. (1987). Nonrandom Loss of Maternal Chromosome-11 Alleles in Wilms- Tumors. *American Journal of Human Genetics* **40**, 413-420.

Seidl, C. I. M., Stricker, S. H. and Barlow, D. P. (2006). The imprinted Air ncRNA is an atypical RNAPII transcript that evades splicing and escapes nuclear export. *Embo Journal* **25**, 3565-3575.

Sharma, P. M., Bowman, M., Madden, S. L., Rauscher, F. J. and Sukumar, S. (1994). RNA Editing in the Wilms-Tumor Susceptibility Gene, Wt1. *Genes & Development* **8**, 720-731.

Sharma, P. M., Yang, X. M., Bowman, M., Roberts, V. and Sukumar, S. (1992). Molecular-Cloning of Rat Wilms-Tumor Complementary-DNA and a Study of Messenger-RNA Expression in the Urogenital System and the Brain. *Cancer Research* **52**, 6407-6412.

Simons, K. and Toomre, D. (2000). Lipid rafts and signal transduction. *Nature Reviews Molecular Cell Biology* **1**, 31-39.

Skarnes, W. C. (2000). Gene trapping methods for the identification and functional analysis of cell surface proteins in mice. In *Applications of Chimeric Genes and Hybrid Proteins, Pt C*, vol. 328, pp. 592-615.

Skryabin, B. V., Sukonina, V., Jordan, U., Lewejohann, L., Sachser, N., Muslimov, I., Tiedge, H. and Brosius, J. (2003). Neuronal untranslated BC1 RNA: Targeted gene elimination in mice. *Molecular and Cellular Biology* **23**, 6435-6441.

Sleutels, F., Zwart, R. and Barlow, D. P. (2002). The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* **415**, 810-813.

Smith, A. G. (1991). Culture and differentiation of embryonic stem cells. *Journal of tissue culture methods* **13**, 89-94.

Smith, T. F. and Waterman, M. S. (1981). Identification of Common Molecular Subsequences. *Journal of Molecular Biology* **147**, 195-197.

Soriano, P. (1999). Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nature Genetics* **21**, 70-71.

Storz, G. (2002). An expanding universe of noncoding RNAs. *Science* **296**, 1260-1263.

Strittmatter, S. M., Fankhauser, C., Huang, P. L., Mashimo, H. and Fishman, M. C. (1995). Neuronal Pathfinding Is Abnormal in Mice Lacking the Neuronal Growth Cone Protein Gap-43. *Cell* **80**, 445-452.

Stumpo, D. J., Bock, C. B., Tuttle, J. S. and Blackshear, P. J. (1995). Marcks Deficiency in Mice Leads to Abnormal Brain-Development and Perinatal Death. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 944-948.

Takasaki, A., Hayashi, N., Matsubara, M., Yamauchi, E. and Taniguchi, H. (1999). Identification of the calmodulin-binding domain of neuron- specific protein kinase C substrate protein CAP-22/NAP-22 - Direct involvement of protein myristoylation in calmodulin- target protein interaction. *Journal of Biological Chemistry* **274**, 11848-11853.

Tatusova, T. A. and Madden, T. L. (1999). BLAST 2 SEQUENCES, a new tool for comparing protein and nucleotide sequences. *Fems Microbiology Letters* **174**, 247-250.

Terashita, A., Funatsu, N., Umeda, M., Shimada, Y., Ohno-Iwashita, Y., Epand, R. M. and Maekawa, S. (2002). Lipid binding activity of a neuron-

specific protein NAP-22 studied in vivo and in vitro. *Journal of Neuroscience Research* **70**, 172-179.

Thomas, K. R. and Capecchi, M. R. (1987). Site-Directed Mutagenesis by Gene Targeting in Mouse Embryo-Derived Stem-Cells. *Cell* **51**, 503-512.

Tufarelli, C. (2006). The silence RNA keeps: cis mechanisms of RNA mediated epigenetic silencing in mammals. *Philosophical Transactions of the Royal Society B-Biological Sciences* **361**, 67-79.

Vergara, G. J., Irwin, M. H., Moffatt, R. J. and Pinkert, C. A. (1997). In vitro fertilization in mice: Strain differences in response to superovulation protocols and effect of cumulus cell removal. *Theriogenology* **47**, 1245-1252.

Vize, P. D., Seufert, D. W., Carroll, T. J. and Wallingford, J. B. (1997). Model systems for the study of kidney development: Use of the pronephros in the analysis of organ induction and patterning. *Developmental Biology* **188**, 189-204.

Wagner, K. D., Wagner, N., Vidal, V. P. I., Schley, G., Wilhelm, D., Schedl, A., Englert, C. and Scholz, H. (2002a). The Wilms' tumor gene Wt1 is required for normal development of the retina. *Embo Journal* **21**, 1398-1405.

Wagner, N., Wagner, K. D., Hammes, A., Kirschner, K. M., Vidal, V. P., Schedl, A. and Scholz, H. (2005). A splice variant of the Wilms' tumour suppressor Wt1 is required for normal development of the olfactory system. *Development* **132**, 1327-1336.

Wagner, N., Wagner, K. D., Schley, G., Coupland, S. E., Heimann, H., Grantyn, R. and Scholz, H. (2002b). The Wilms' tumor suppressor Wt1 is associated with the differentiation of retinoblastoma cells. *Cell Growth & Differentiation* **13**, 297-305.

Wang, W. H., Duan, J. X., Vu, T. H. and Hoffman, A. R. (1996). Increased expression of the insulin-like growth factor-II gene in Wilms' tumor is not dependent on loss of genomic imprinting or loss of heterozygosity. *Journal of Biological Chemistry* **271**, 27863-27870.

Wang, Z. Y., Qiu, Q. Q. and Deuel, T. F. (1993). The Wilms-Tumor Gene-Product Wt1 Activates or Suppresses Transcription through Separate Functional Domains. *Journal of Biological Chemistry* **268**, 9172-9175.

Wang, Z. Y., Qiu, Q. Q., Gurrieri, M., Huang, J. and Deuel, T. F. (1995). Wt1, the Wilms-Tumor Suppressor Gene-Product, Represses Transcription through an Interactive Nuclear-Protein. *Oncogene* **10**, 1243-1247.

Ward, A. (1997). Beckwith-Wiedemann syndrome and Wilms' tumour. *Molecular Human Reproduction* **3**, 157-168.

Widmer, F. and Caroni, P. (1990). Identification, Localization, and Primary Structure of Cap-23, a Particle-Bound Cytosolic Protein of Early Development. *Journal of Cell Biology* **111**, 3035-3047.

Wiederkehr, A., Staple, J. and Caroni, P. (1997). The motility-associated proteins GAP-43, MARCKS, and CAP-23 share unique targeting and surface activity-inducing properties. *Experimental Cell Research* **236**, 103-116.

Wilhelm, D. and Englert, C. (2002). The Wilms tumor suppressor WT1 regulates early gonad development by activation of Sf1. *Genes & Development* **16**, 1839-1851.

Wroe, S. F., Kelsey, C., Skinner, J. A., Bodle, D., Ball, S. T., Beechey, C. V., Peters, J. and Williamson, C. M. (2000). An imprinted transcript, antisense to Nesp, adds complexity to the cluster of imprinted genes at the mouse Gnas locus. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 3342-3346.

Yagi, T., Ikawa, Y., Yoshida, K., Shigetani, Y., Takeda, N., Mabuchi, I., Yamamoto, T. and Aizawa, S. (1990). Homologous Recombination at C-Fyn Locus of Mouse Embryonic Stem-Cells with Use of Diphtheria-Toxin a-Fragment Gene in Negative Selection. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 9918-9922.

Yamamoto, Y., Sokawa, Y. and Maekawa, S. (1997). Biochemical evidence for the presence of NAP-22, a novel acidic calmodulin binding protein, in the synaptic vesicles of rat brain. *Neuroscience Letters* **224**, 127-130.

Ye, Y., Raychaudhuri, B., Gurney, A., Campbell, C. E. and Williams, B. R. G. (1996). Regulation of WT1 by phosphorylation: Inhibition of DNA binding, alteration of transcriptional activity and cellular translocation. *Embo Journal* **15**, 5606-5615.

Yeger, H., Cullinane, C., Flenniken, A., Chiltonmacneill, S., Campbell, C., Huang, A., Bonetta, L., Coppes, M. J., Thorner, P. and Williams, B. R. G. (1992). Coordinate Expression of Wilms-Tumor Genes Correlates with Wilms-Tumor Phenotypes. *Cell Growth & Differentiation* **3**, 855-864.

Yelin, R., Dahary, D., Sorek, R., Levanon, E. Y., Goldstein, O., Shoshan, A., Diber, A., Biton, S., Tamir, Y., Khosravi, R. et al. (2003). Widespread occurrence of antisense transcription in the human genome. *Nature Biotechnology* **21**, 379-386.

Young, T. L., Matsuda, T. and Cepko, C. L. (2005). The noncoding RNA Taurine upregulated gene 1 is required for differentiation of the murine retina. *Current Biology* **15**, 501-512.

Zakharov, V. V., Capony, J. P., Derancourt, J., Kropolova, E. S., Novitskaya, V. A., Bogdanova, M. N. and Mosevitsky, M. I. (2003). Natural N-terminal fragments of brain abundant myristoylated protein BASP1. *Biochimica Et Biophysica Acta-General Subjects* **1622**, 14-19.